View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by University of Houston Institutional Repository (UHIR)

ESSAYS ON THE ROLE OF SPECULATION IN THE VOLATILITY OF OIL PRICES AND OIL FUTURES RISK PREMIA

A Dissertation

Presented to

The Faculty of the Department

of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Mahboobeh Asghari

May 2015

ESSAYS ON THE ROLE OF SPECULATION IN THE VOLATILITY OF OIL PRICES AND OIL FUTURES RISK PREMIA

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Mahboobeh Asghari

May 2015

Abstract

The greater volatility and higher prices in the oil market after 2003 suggest the possibility of increased speculation. Many researchers have investigated the origin of this change in oil price dynamics. The literature, however, lacks a comprehensive empirical model of oil price dynamics that incorporates both fundamentals and the influence of all economic agents who behave speculatively, and is also able to distinguish between speculation which is necessary for the oil market and "excess speculation".

The first chapter introduces a comprehensive empirical model which incorporates all of these elements, and uses it to examine the contribution of fundamentals and excess speculation to oil price fluctuations. I do this by introducing three novel features to the literature. First, I introduce an operational definition for excess speculation. Second, using a dynamic common factor model, I extract orthogonal factors from monthly time-series data that are specific to economic fundamentals and purely speculative activities, respectively, and decompose the variance of oil prices using these orthogonal factors. Third, I employ a larger set of underlying variables to represent speculative activities. My results show that the contribution of excess speculation in this market increased to 28% after 2003 compared to 1% before 2003, which shows a structural break in speculators' contribution after 2003.

Moreover, one of the six strands of the literature considers variation in oil risk premia

to address the existence of speculative activities in oil futures market, and a structural change in oil futures markets in recent years. The literature, however, lacks a general model of time-varying risk premia that incorporates all of the possible explanatory variables of oil price fluctuations.

The second chapter employs this approach, time-varying oil risk premia, to investigate the role of speculative activity in this market, by introducing two novel features to the literature. First, I consider the notion of time-varying risk premia along with other explanations of oil price fluctuations. Second, I provide evidence of excess speculation based on these models. My results verify the variation of oil futures risk premia in recent years.

Acknowledgements

I would like to thank Dietrich Vollrath, Bent Sorensen, and Christian Murray for all of their advice, help, and support.

Specifically, I would like to thank Dietrich Vollrath for being a great advisor, Bent Sorensen for being a considerate graduate director, David Papell for being a supportive chair, Amber Pozo for all of her help during these five years, Anita Gaines and all other kind staffs in International Student and Scholar Services (ISSSO) for smoothing the education path for international students like me, and all my classmates for making this path enjoyable.

To my wonderful father, mother, brother, and sister

whom I missed every second of these years...

Contents

List of Tables

List of Figures

Chapter 1

Did Speculation Cause Greater Volatility in Oil Prices?

1.1 Introduction

Starting in about 2003, a surge in both the volatility and price level of oil prices has captured attention from academia and the public, and caused a public debate about the determinants of the price of crude oil. Many researchers have tried to find if this surge can be explained by economic fundamentals, or if it is attributable to speculators' activity in the oil futures market. This has led to the existence of two competing points of view about the major determinants of oil prices in the literature: one relates it to changes in fundamentals, and the other attributes it to increased speculators' activity as a result of financialization of oil markets.

This paper develops a new methodology to answer the following questions. First, what is the major determinant of oil price fluctuations, fundamentals or "excess speculation"? Second, how have the contribution of these two sources evolved since 2003? The answers to these questions have implications for evaluating the two competing views about the major determinant of oil price fluctuations, and for policymakers who have concerns about regulating this market.

My main finding is that excess speculation is the major driver of oil price fluctuations after 2003. This is a robust feature of the data, and it is not simply an effect of the higher volatility in oil price fluctuations in this period. When considering monthly frequency data for the whole period from 1986 to 2013, only 3.4 percent of oil price fluctuations can be explained by fundamentals, while 2 percent of oil price fluctuations are attributable to excess speculative activities. But after 2003, the contribution of real activities to oil price fluctuation increases to 9.3 percent, while the contribution of excess speculation increases to 27.5 percent. The remaining contribution is attributable to idiosyncratic dynamics in oil prices. There has been a decline over time in the relative importance of fundamental factors in accounting for oil price fluctuations. In other words, the effect of speculative activities in driving the oil price is substantially increasing over time, especially during the period of high volatility in oil price fluctuations.

There are three main reasons why distinguishing between these two points of view, and identifying the role of speculation in the oil market is difficult. This paper attempts to make progress on all three:

First, there were dramatic changes in the world economy which affected the demand and supply of oil at the same time that oil prices became more volatile. The development of emerging economies such as China and India generated higher demand for oil. There was also increased uncertainty on the supply side, such as below projection supplies by OPEC and non-OPEC countries, and political risk in oil producing countries such as Russia.[1] As a result, some researchers (Kilian and Murphy, 2014) believe that the surge in prices and volatility in the oil market originates with this dramatic change in fundamentals.

At the same time, there were dramatic changes in the financialization of the oil market, including technological advances in oil futures markets which gave financial investors such as hedge funds easy access to oil derivatives, resulting in an influx of investors to this market seeking high returns. As early as 2004, policy makers such as Ben Bernanke were concerned, and warned about the appearance of this possibly speculative component in oil prices. Michael W. Masters (2008), in testimony before the U.S. Senate Committee on Homeland Security and Governmental Affairs, unequivocally attributes the new trend in oil volatility to the activity of a new type of investors in this market, which he referred to as "index speculators". He asserted

[1] See comments by former Federal Reserve chairman, Ben Bernanke: www.federalreserve.gov/boarddocs/speeches/2004/20041021/default.htm

that the index speculators, unlike the traditional speculators, do not buy futures and take delivery, but rather they roll their entire long position forward and need never close these positions. As a result, their activity leads oil prices to be more volatile.

The occurrence of these two dramatic changes can each be considered as a potential explanation for the new dynamics in oil prices, making it difficult to identify the empirical role of each separately.

Second, from an empirical perspective it is not always obvious how to measure speculation. The broad definition of speculation from Killian and Murphy (2011) is that a speculator is "anyone buying crude oil not for current consumption, but for future use" in the hope of profit in the future. In this case, speculation is not necessarily attributable to pure speculators, but also includes oil producers and oil companies reducing current extraction (inventory under the ground) and keeping oil in inventory above ground, in the pursuit of selling the oil at a higher price in the future. Hence, there are at least two broad parties that behave speculatively according to Killian and Murphy: pure speculators and oil companies. It is important to include the speculation activity of all of these parties when gauging the extent of speculation, but also to differentiate between the two.

Third, speculation (by pure speculators or oil companies) may or may not be closely related to fundamentals. If there are expectations of rising oil prices, as in anticipation of a disruption of oil supplies or changes in oil demand, speculators increase oil futures prices by buying oil futures contracts, which signal high spot prices in the future. As a result, there is higher demand for oil inventory, both extracted and unextracted. This, in turn, leads to an increase in current oil prices, a decrease in current oil consumption, and an incentive for oil producers to increase their capacity of production in the future. In this manner, speculation provides liquidity to oil markets and helps the functions of price discovery and risk transfer between the future and the present. This speculation is linked to changes in fundamentals and, hence, as discussed by Fattouth et al. (2012), there may be no point in preventing this kind of speculation. On the other hand, "excess speculation" may prevent the oil market from functioning properly, generating changes in the dynamic of oil prices unrelated to fundamentals. This speculation is only favorable from a private perspective, and is not necessary for the market to react to fundamental shocks. Therefore, it is critical to distinguish between these two types of speculative activity, but in practice making this empirical distinction is difficult.

Because of these issues, there is no clear definition on the role of speculation in oil price movements. As discussed in Fattouth et al. (2012), "the academic literature has focused on the problem of finding indirect evidence in support of the view that financial speculation has affected oil prices" and "no one has been able to quantify the problem of excessive speculation to date. Rather the public debate has tended to conflate several recent developments in oil futures markets and in the spot market for crude oil, taking the existence of excessive speculation as self-evident."

This paper uses a new methodology to examine the contribution of fundamentals and excess speculation to oil price fluctuations, and analyzes if the contribution of fundamentals and excess speculation to oil price fluctuations has evolved over the period 1986-2013. The basic idea in this paper is as follows. Assume that there is one index which is representative of real activity in the economy ("fundamentals"), and there is another index which captures all speculative activities by both pure speculators and oil companies. The part of this total speculative activity which cannot be explained by expectations of fundamentals is thus "excess speculation", and is found as the residual from a regression of the speculative index on the expectation of the fundamentals index. Excess speculation is measured as the part of all speculative activity that is orthogonal to fundamentals. Regressing oil price fluctuations on the original fundamental index and the residual excess speculative index allows me to calculate how much of the variation in oil prices is caused by each.

In practice, I use a dynamic factor model (DFM) instead of the simple "indexes" mentioned above to capture fundamentals and overall speculative activity. I then find the part of overall speculative activity orthogonal to fundamentals, and label this "excess speculation". The dynamic factor model allows me to capture the common component of fluctuations in a large set of variables. The key output is a set of latent dynamic factors that summarize the common movement behind a high-dimensional vector of time-series data. The dynamic factor model does this without having to make strong identifying assumptions or imposing a prior distribution on the parameters.

Fundamentals are therefore captured by a vector of latent factors common in a set of up to 24 real variables, such as industrial production and housing starts. I find the common factors of overall speculative activity from a set of up to 11 variables related to speculative activity, such as a proxy for speculators' risk-aversion. The excess speculation common factor can then be found by regressing the overall speculative factors on the fundamental factors and taking the residuals, calling these the excess speculation factors.

For my method, deciding what variables to include in the dynamic factor model (DFM) is crucial. Variables which can gauge the extent of uncertainty in the oil market provide me with the pool of potential explanatory variables for oil price fluctuations. According to economic theory, similar economic forces drive both oil futures risk premia and oil price fluctuations.[2] The futures risk premium reflects the compensation for the uncertainty perceived in the oil market. Therefore, changes in uncertainty provide information on changes in oil prices. I classify the variables measuring uncertainty into two groups, fundamental and speculative variables.

One source of uncertainty is in the situation of the whole economy, and this is measured by fundamental variables. The second source of uncertainty is specific to the oil market, and can be measured by uncertainty in the demand and supply of oil. Variables measuring this second source of uncertainty, capture the overall activity of both groups of speculators, oil producers and pure speculators.

My paper contributes to the literature in three dimensions: First and most importantly, I provide a clear empirical definition of excess speculation. The methodology that I implement defines excess speculation factors as the component of all speculative activity that is orthogonal to fundamental factors. Using these two orthogonal factors, I can distinguish clearly between oil price movements driven by fundamentals versus excess speculation.

Second, I make use of a much larger set of variables than the prior literature to capture fundamental and speculative activities. This allows me to capture thoroughly the commonality in each category of fundamental and speculative activities. Further, the long time horizon enables me to consider two distinct sub-periods and, in particular, enables me to analyze how the contribution of fundamentals and speculation to oil price fluctuations has evolved during the period of sharp increases in volatility after 2003.

[2] See Alquist and Kilian (2010) and Acharya et al. (2013).

Third, I employ a set of recently developed econometric tools to analyze the question of interest. Dynamic factor model captures the commonality of fluctuations in either fundamentals or speculative variables without making strong identifying assumptions or imposing a prior distribution on the parameters. In comparison, the prior literature which uses structural VAR requires identifying assumptions which are somewhat arbitrary and not generally compelling (Hamilton, 1994).

The results in this paper have important implications for policy makers who have concerns about regulating the oil market such that they ensure this market functions properly and the excess speculation activity no longer impacts this market. There may be no point to impose restriction on the speculators' activity in the oil market if it merely reacts to changes in the real economy. The results in this paper, however, lay out evidence of "excess speculation" in this market, which originates from pure speculators' activity especially after 2003, and it is not a function of fundamentals.

The reminder of the paper is as follows. Section 1.2 reviews the existing literature on the role of speculation and on commodity markets. Section 1.3 provides an underpinning theory about how I classify the variables into two groups of fundamentals and speculation. Section 1.4 discusses the empirical methodology of the paper in detail. Section 1.5 mentions the data sources. Section 1.6 lays out the results, and finally, Section 1.7 concludes.

1.2 Literature Review

1.2.1 Speculation, Excess Speculation and Findings on Speculation in Oil Market

Investigating the role of speculation is difficult. There are at least two broad parties of investors in the oil market who behave speculatively: pure speculators and oil companies. These two parties are not isolated from each other and the behavior of one party of speculators can affect the behavior of the other group while both groups are trying to maximize profits. Defining the extent of speculation requires first quantifying the speculation by all of these parties.

Moreover, speculation can be classified as desirable speculation or excessive speculation. Desirable speculation provides liquidity to the commodity market and helps price discovery and risk transfer, while excess speculation impacts the dynamics of commodity prices adversely. For instance, it might increase the price volatility, and, as a result of the increased uncertainty about energy costs, companies might make less capital investments leading to slack in the economy. Therefore, it is necessary to distinguish between desirable speculation and excessive speculation.

Fattouh et al. (2012) identify six strands in the literature which investigate the role of speculation along with the critiques about why the speculation identified is not necessarily compelling. In this section, I first discuss the six strands and their critiques from Fattouh et al. (2012). Then, I will offer my own critiques of the existing literature.

First strand. The first strand addresses the increased comovement among the spot price of oil, oil futures prices, non-energy commodity prices, foreign exchange, and stock prices, or the higher correlation among their returns, as a result of financialization of oil futures markets after 2003. It then argues that the new dynamics in oil prices were caused by the activity of new financial investors in this market (See e.g. Tang and Xiong 2010).

The first critique is that, according to economic theories, changes in global fundamentals will be reflected in both spot and futures prices, leading to a comovement between them. Thus examining the comovement between spot and futures prices to identify financialization might be invalid. The second critique is the surge in the price of non-exchange traded commodities after 2003, making it hard to consider the financialization in the commodity market as a potential explanation for this new trend in prices.

Second strand. Another strand investigates if the increased flows to oil index funds predict changes in oil prices, which may be interpreted as the influence of financial investors on price dynamics in this market (Singleton 2012).

The first critique on this strand is that including precise data on individual traders' positions (such as hedge funds' positions) rather than aggregated CFTC data on all investors' positions, leads to different conclusions in precedence of predictability. For instance, Büyüksahin and Harris (2011) find futures price changes precede changes in positions of hedge funds or other non-commercial investors, which then implies speculators respond to the real economy. Which one, flows to index funds or the changes in oil futures prices, precedes the other one may lead to different arguments about the role of speculation.

The second critique is that even evidence that investors' flows precede price movement does not necessarily show the speculation. In the case that the activity of non-commercial traders in the futures market is limited because of informational costs, an increase in hedging demand increases the risk premium, which will be reflected in oil futures returns and spot price fluctuations, while higher hedging demand may originate from the expectation of higher real economic activity.

Third strand. The third strand examines if there is a casual relation between futures prices and the spot price, and argues that higher futures prices cause increases in the spot price of oil, as the oil futures price is used as a benchmark.

The critique is that the same economic forces drive both oil futures prices and spot prices. It might be the case that expectations about oil demand and supply or real economic activity increase oil futures prices, and then, according to economic theories, these expectations would be reflected in the oil spot price. Hence, finding a link between futures prices and the spot price cannot be taken as evidence of speculation

Fourth strand. Another strand considers the evolution of oil inventories to address the existence of speculation. In recent years, despite high oil inventories, oil prices have increased. In other words, there is a positive relation between oil inventories and oil prices after 2003, rather than a negative relation, which might be a sign of speculation in this market.

The first critique is that the positive relation between the oil futures spread (which might be considered as a speculative component of the real oil price according to Fattouh et al. (2012)) and the change in world oil inventories may imply speculation, while this strand considers the positive relation between nominal oil prices and U.S. inventories, which is misleading for investigating the role of speculators.

Second, even finding a positive relationship between the oil futures spread and the change in world oil inventories can indicate the role of speculation if one assumes that the speculators' activity exogenously increases oil futures prices, which in turn lead to an increase in oil inventories. In the case that the speculators behave endogenously in response to the real economy, the increase in oil inventories cannot be taken as an evidence of speculation. For example, expectations about higher oil demand or shortages in oil supply may encourage more buyers of oil futures contracts and higher futures prices.

Fifth strand. A fifth strand considers models in which the risk premium in the commodity market (compensation on the risk related to a futures contract) evolves over time, and takes it as evidence of speculative activities in oil futures market. This strand also considers large changes in the term structure of oil futures and the shift in the front end and the back end of the term structure of oil futures as indicators of a structural change in oil futures markets in recent years. I discuss modeling the time-varying risk premia in more details in chapter 2.

The critique is that although time-varying risk premia might be suggestive of the influence of financial investors on oil futures markets, there is no evidence of excess speculation based on these models. It is also not obvious which of time-varying risk premia or structural VAR has better explanatory power and predictability of oil price fluctuations.

Sixth strand. The last strand of the literature is one which uses structural vector autoregression models of the oil market to examine speculation. Structural VARs allow a disentangling of the different shocks to oil price fluctuations, including oil supply and oil demand shocks as well as speculative shocks, and gauging the effect of each shock separately on oil price fluctuations.

Kilian and Murphy (2011) were the first structural VAR in this literature and included four variables: global oil production, a measure of global real activity, the real price of crude oil, and the change in above-ground global crude oil inventories. They reason that shocks to above-ground global oil inventories can be interpreted as speculative shocks when the price elasticity of oil demand is different from zero. As Hamilton (2009b) discusses, if the price elasticity is not zero, then a change in oil inventories must signal speculation. In this case, an increase in oil prices as a result of speculators' activity in the oil futures market leads to a reduction in oil consumption and in turn, a reflection in oil inventories. Kilian and Murphy (2011) show that the short-run price elasticity of oil demand is about -0.25. Thus, they argue any increase in oil prices as a result of speculators' activity in the oil futures market is reflected in involuntary accumulation of oil inventories. They find changes in oil prices after 2003 can be well explained by positive shocks to world oil demand through global real activity. In contrast, the surge in oil prices is not explained by speculative demand for inventories. Their results show speculation by oil producers or other negative oil supply shocks had little impact on the real price of oil.

In two other studies, Juvenal and Petrella (2011) and Lombardi and Van Robays (2011) argue that the Kilian and Murphy model does not capture "financial speculation". Hence, in a structural VAR model, they introduce other speculative shocks besides above-ground global oil inventories shock in Kilian and Murphy (2011). Juvenal and Petrella (2011) consider under-ground oil inventory shocks as speculative supply shocks by oil producers, and Lombardi and Van Robays (2011) define speculative shocks as a shock that leads to a contemporaneous increase in both the oil futures spread and the oil futures price.

As discussed by Fattouh et al. (2012), the speculative supply shock in Juvenal and Petrella (2011) can be considered the same as the flow supply shock in Kilian and Murphy (2011) since it impacts the oil supply. Moreover, standard models of storage assert that a speculative demand shock results in a negative oil futures spread, while Lombardi and Van Robays (2011) consider the positive oil futures spread as a sign of speculative activity.

Another paper that investigates the role of speculation is Knittle et al. (2013). Their work can be classified in the fourth strand (which involves inventories) discussed above, where they apply a new method for defining the contribution of speculation. They express the oil price as a function of oil supply and demand as well as oil inventories. Then, they use a variance decomposition to identify that part of oil price fluctuations due to inventory changes, holding supply and demand contributions constant. They attribute this fraction of oil price fluctuations to speculation over the 1999-2012 period and conclude that it is trivial. The results are questionable. First, the effect of the change in oil demand and supply only reflects in the shift in oil demand and supply while other parameters such as elasticities of oil supply and demand are assumed to be constant implying no structural break in the oil market. Second, they argue that the shift in demand for storage is due solely to speculative activity, ruling out a role of fundamentals for the shift.

1.2.2 Critiques of the Existing Literature

Regardless, the advantage of the structural VAR is that it incorporates a role for changes in economic fundamentals. In comparison, as discussed above, by ignoring real activity, all of the evidence on speculation provided in the other five strands discussed by Fattouh et al. (2012) might be explained by changes in real economic activity.

Fattouh et al.(2012) assert that the structural VAR in Killian and Murphy (2011) "provides the most decisive and most formal evidence on the question of speculation to date." In that paper, the interpretation of oil inventory shocks as speculative shocks is based on the assumption of unlimited arbitrage between the oil futures market and the oil spot market.[3]

There are several caveats regarding Killian and Murphy (2011, 2014) interpretation, as follows. First, the assumption of unlimited arbitrage itself is questionable. There is a limited capacity for holding oil above ground, and the response of oil inventories to contemporaneous changes in the real economy and speculation in oil futures market is constrained. In fact, Killian and Murphy (2011, 2014) implicitly assume that the elasticity of storage space to inventory demand is positive, but this seems unlikely to hold in the short run. As such, by ordering oil inventories last in a SVAR, the structural shock to oil inventories, which they identify as the speculative component, can be made arbitrarily small. This identification assumption invalidates the argument that considering other speculative variables such as "the oil futures spread are redundant" in their model.

Second, they do not allow for time-varying volatility in structural inventories shocks or a structural break in volatility, which might invalidate the results. They imply that "percentage change in inventories does not appear to be covariance stationary, whereas the change in inventories does". As a result, they include the change in inventories in the SVAR model to resolve this problem, which is not compelling. Models of time-varying risk premia suggest there is a structural break in the oil market and

[3] They also specify the real oil price in log levels, assuming its stationary while I believe that this is an invalid assumption. The question of oil prices' stationarity is beyond the scope of this paper.

the interaction of oil producers and speculators is not constant over time. In the case that we allow for time-variation in risk premia, the results in their paper are no longer reliable.

Third, if speculators' activity increases the expected default frequencies of oil companies, oil inventories cannot be chosen as a proper proxy for capturing the speculation in oil markets. As Acharya et al. (2013) demonstrate, the reaction of oil companies to an increase in expected default frequency would be a reduction in oil inventories.

In my paper, I address the caveats mentioned above. First, I incorporate other variables rather than oil inventories to capture speculative activities. Second, I employ variables which impact time-varying risk premia as potential explanatory variables of oil price fluctuations. Both spot prices and futures prices respond to the same economic forces according to economic theory, hence variables which gauge risk premia can be employed as the explanatory variables of oil price fluctuations. The advantage is that by including these variables, I allow for evolution of the risk premia and as a result, the evolution of oil price dynamics over time, rather than assuming the constant dynamics which might lead to invalid results. Third, I construct an empirical model which captures the interaction between oil producers and speculators.

Moreover, I incorporate two groups of orthogonal factors which enables me to disentangle the different shocks to oil prices, which is essential for an accurate answer to this question. I also employ a common factor method which allows me to consider the contribution of the different sources to oil price fluctuations in a flexible manner, rather than making some arbitrary identification assumptions in a SVAR, which are not necessarily consistent with economic theories.

Thus, I construct an empirical model for investigating oil price fluctuations based on time-varying risk premia which allow the time variation in interaction between oil producers and speculators. The methodology employed also allows me to address the existence of excess speculation if there is any.

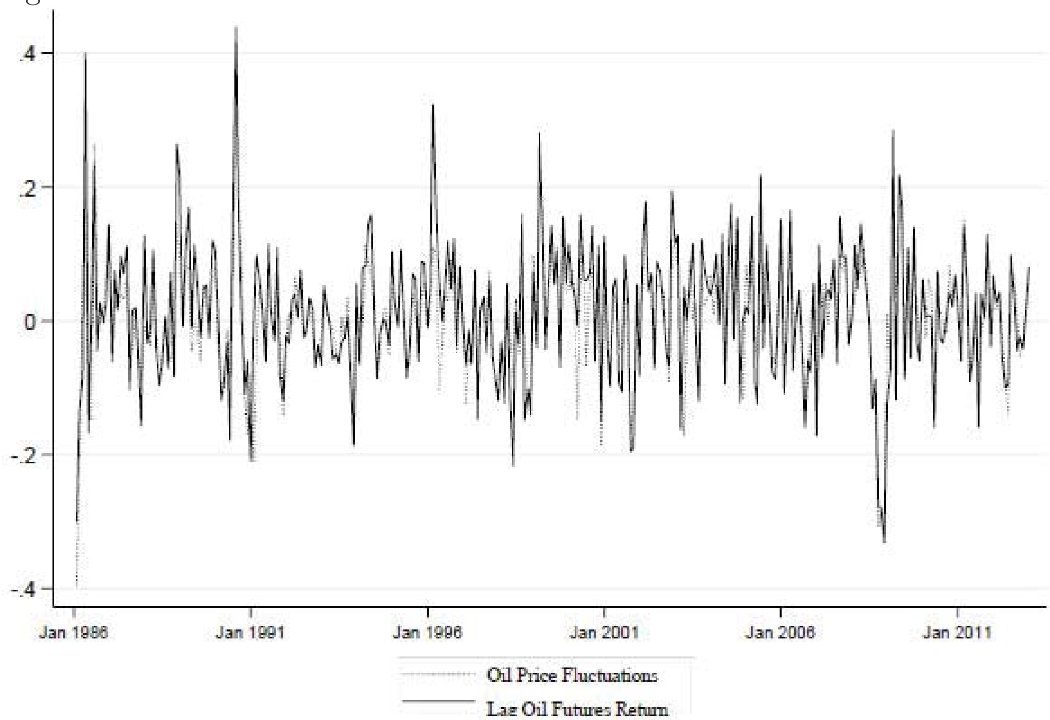
1.3 The Theoretical Underpinning for the Variables in the Empirical Methodology

Deciding what variables to include in the dynamic factor model (DFM) as fundamental and speculative variables is crucial. The business cycle variables can be considered as fundamental variables. The speculative variables can be considered as variables that distort oil futures risk premia from the risk-adjustment of equity holders and capture the impact and interaction of two groups of speculators (pure speculators and oil producers) on oil futures risk premia. Below I explain the logic behind this.

Potential explanatory variables to explain oil price fluctuations, in this paper, are chosen based on the model of time-varying risk premia in Acharya, Lochstoer and Ramadorai (2013). The variables employed in their paper to measure the extent of uncertainty and, as a result, for measuring the oil futures risk premium, provide me with a means for determining the sources of oil price fluctuations.

It is important to emphasize that according to economic theory, any changes in oil futures risk premia must be reflected in fluctuation in oil spot prices by construction. In fact, the same economic forces drive both spot prices and futures prices. Figure 1.1 shows the comovement between oil spot price fluctuations and oil futures risk premia. Thus, variables which gauge oil futures risk premia can be employed for explaining fluctuation in oil spot prices.

Figure 1.1: Comovement Between Oil Price Fluctuations and Oil Futures Returns



The model in Acharya et al. (2013) provides the theoretical underpinning for the variables included and the empirical strategy introduced in this paper, and can also be helpful in understanding the empirical methodology section. Thus, I briefly explain some aspects of this model.

Acharya et al. (2013) introduce a two-period general equilibrium model of commodity spot and futures price determination. Comparative statics for commodity spot and futures prices with regard to the variables are then derived from the reduced form of the model. The three types of agents in the model are as follows.

Consumers who are demanders of spot commodity.

Commodity producers are price takers and have risk-averse managers. Managers face a tradeoff between profit and earnings variance in the next period. This assumption induces the risk-averse manager to optimally manage the inventory and to optimally hedge with commodity futures. When hedging, producers hold more inventory and in turn, they decrease the risk related to next period's earning. If the futures price is lower than the expected future prices, the producer can increase expected profits by entering a long speculative futures position. An increase in the manager's risk aversion decreases this speculative futures position.

Speculators take the long positions, rather than the short positions taken by producers, and the interaction of these two groups determines the futures price. On the other hand, speculators are risk-neutral, and the speculators face capital constraints due to leverage costs or Value-at-Risk (VaR) limitations.

In the stochastic general equilibrium (SGE) framework, in which each agent maximises its objective function subject to a constraint it faces and then a general equilibrium is considered, they derive the futures risk premium (the compensation for the uncertainty in the oil market). In the case that the speculators do not face the capital constraint, the futures risk premium would be equal to the excess market return. Incorporating the capital constraint in the model, the futures risk premium in this market is no longer equal to excess market return (equity risk premium), which can be gauged by business cycle variables. Rather, the futures risk premium also depends on the risk-aversion of speculators and producers. Specifically, the futures risk premium and the expected percentage point change in the spot price are increasing in producers' risk-aversion (producers' fundamental hedging demand) and speculators' risk-aversion (the degree of speculators' capital constraints).

The intuition for the effect of producers' risk-aversion and speculators' risk-aversion on the futures risk premium and oil price fluctuations is as follows. An increase in producers' risk-aversion causes the producers to be more cautious about holding unhedged inventory. As a result, they increase the number of short futures positions inducing a higher variance-adjusted demand in the futures market. The producers also decrease oil inventory leading the current spot prices to decrease and future spot price to increase. Overall, these two strategies result in the higher futures risk premium and expectations for higher spot price fluctuations. Moreover, an increase in risk-aversion of speculators means that there should be more compensation for the risk they are tolerating. The producers respond to this by decreasing the number of short futures positions (lower short hedger positions), which leads to having more unhedged inventory and being at more risk in the next period. The producers in turn reduce their inventory which depresses the current spot price. The overall effect is higher futures risk premium and the higher expected spot price fluctuations.

In addition, the extent to which fundamental hedging demand impacts the futures risk premium depends on the variance of the commodity future returns as well as speculator risk tolerance. In particular, the impact of fundamental hedging demand on the futures risk premium is higher when the variance of the commodity futures returns is high (the effect can be captured by an interaction term between fundamental hedging demand and the variance of the commodity futures returns, as shown in the empirical section). Also, a change in fundamental hedging demand has a larger effect on the futures risk premium if speculator risk tolerance is low, which requires greater compensation reflected in the higher futures risk premium (the effect can be captured by an interaction term between fundamental hedging demand and speculator risk tolerance, as shown in the empirical section). Indeed, the channel that the speculators' risk-aversion impacts futures and spot commodity prices is through the hedging demand of producers.

Thus, the variables included in their paper gauge time-varying oil futures risk premia. The futures risk premia in the oil market reflects the compensation for two sources of uncertainty in the oil futures market. Therefore, I require variables which gauge these two sources.

The first source of variation in risk premia in the oil market is associated with the uncertainty in the whole economy and can be gauged by business cycle variables. This is represented by the excess market return. Thus, business cycle variables can be employed as a forecaster of excess market returns, which are related to the conditional covariance of futures returns with equity holders' pricing kernel. For instance, Acharya et al. (2013) use business cycle variables such as the default spread (the difference between the Baa and Aaa rated corporate bond yields, which captures aggregate default risk in the whole economy) as a forecaster of excess returns on stocks and bonds[4] , and the risk-free rate (such as 3M T-bill rate) as a robust, out-of-sample predictor of equity market excess returns.[5]

The second source of variation in risk premia in the oil market (futures excess return) represents the compensation to oil producers and speculators as the investors in this market for tolerating the idiosyncratic risk specific to the oil market. The futures excess returns can be related to hedging demand (see e.g. De Roon, Nijman, and Veld, 2000). Futures excess returns can also be gauged by the interaction of hedgers and arbitrageurs (see e.g. Hirshleifer, 1988, 1990 and Acharya et al., 2013).

[4]As shown by Fama and French (1989)
[5]As shown by Ang and Bekaert (2007)

1.4 Empirical Methodology

The empirical methodology that I employ requires identification of two different sources of oil price fluctuations, and accounts for the dynamic relationship between the set of fundamentals and speculation variables. In order to understand my approach, I first describe the methodology briefly, and then discuss dynamic factor models following Stock and Watson (2010). Finally I describe the exact nature of my empirical methodology.

1.4.1 Brief Description of the Empirical Methodology

I construct a model that contains two orthogonal groups of factors: (1) fundamental factors common to real variables (F_{fund}), and (2) excess speculation factors common only to speculative variables (F_{exspec}). Letting ΔOil_t denote the growth rate of the oil price, the model can be written as:

$$\Delta Oil_t = \beta_f F_{fund,t} + \beta_s F_{exspec,t} + e_t. \tag{1.1}$$

The coefficients β_f and β_s quantify the extent to which oil price fluctuations move with the fundamental factors and the excess speculation factors.

I use a variance decomposition to assess the relative contributions of fundamental and speculative variables to oil price fluctuations. With orthogonal factors, the variance of oil price fluctuations can be written as follows:

$$var(\Delta Oil_t) = \beta_f^2 var(F_{fund,t}) + \beta_s^2 var(F_{exspec,t}) + var(e_t). \tag{1.2}$$

Then, the fraction of volatility in oil price fluctuations due to changes in fundamentals is

$$\frac{\beta_f^2 var(F_{fund,t})}{var(\Delta Oil_t)}. \tag{1.3}$$

and the fraction of volatility in oil price fluctuations purely due to excess speculative activities in the economy is

$$\frac{\beta_s^2 var(F_{exspec,t})}{var(\Delta Oil_t)}. \tag{1.4}$$

The main empirical issue I have is that there is no single index of "fundamentals", F_{fund}, or of speculative activity, F_{exspec}. Therefore, my methodology uses dynamic factor models to find a vector F_{fund} of the common factors driving fluctuations in real activity, as captured by a set of variables described in section 1.4.3. Similarly, I use dynamic factor models to find a vector F_{exspec} that are the common factors of a set of variables associated with speculative activity in the rest of the economy.

1.4.2 Common Factor

Dynamic factor models ($DFMs$) are time-series extension of factor models initially developed for cross-sectional data. The key finding is that the common movement behind a high-dimensional vector of macroeconomic data can be captured and summarized by a few unobservable (latent) dynamic factors without making strong identifying assumptions or imposing a prior distribution on parameters to separately identify the effects of different shocks. This common trend can explain a large fraction of fluctuations (variance) of each series and is modeled as a vector autoregression (VAR). Besides the common trend, the process of each series also includes a mean-zero idiosyncratic trend which represents the unique feature of that series. Formally, the dynamic factor model can be specified as

$$X_t = \lambda(L) f_t + e_t \tag{1.5}$$

$$f_t = \psi(L) f_{t-1} + \eta_t \tag{1.6}$$

where X_t is a $(N \times 1)$ vector of observable variables (X_t is, for example, a vector of either fundamental or speculative variables), there are q dynamic factors (f_t) and $\lambda(L)$ shows the dynamic factor loading for each series. By assumption, all the processes in (1.5) and (1.6) are stationary, the idiosyncratic disturbances (e_t) are uncorrelated with the factor innovations (η_t) at all leads and lags, and the idiosyncratic disturbances are mutually uncorrelated at all leads and lags.

Equivalently, the dynamic factor model can be rewritten as a linear state space model:

$$X_t = \Lambda F_t + e_t \tag{1.7}$$

$$\Phi(L) F_t = G \eta_t \tag{1.8}$$

which is referred to the "static form" of the DFM, because the factors in this representation appear only contemporaneously, although the static factors include current and past values of the dynamic factors. $F_t = (f_t^{'}, f_{t-1}^{'}, ..., f_{t-p}^{'})'$ is an $(r \times 1)$ vector of r static factors, $\Lambda = (\lambda_0, \lambda_1, ..., \lambda_p)$ where λ_i is the $(N \times q)$ matrix of coefficients on the i^{th} lag in $\lambda(L)$, $\Phi(L)$ is the matrix of $1^{'}$s , $0^{'}$s and the elements of $\psi(L)$, and G is a matrix of $1^{'}$s and $0^{'}$ chosen in a way that the dynamic and static forms of the DFM are equivalent.

There are three ways to specify the time-series extension of DFMs, including: (1) the Kalman filter as a parametric model for estimating the factors between low-dimensional vector of observations; (2) principal components, which is a nonparametric model, for estimating the factors between high-dimensional vector of observations and; (3) hybrid principal components and state space methods.

I cannot use the Kalman filter since I am employing a large set of variables. I also cannot employ the principal components models since there is no dynamic time-series process considered in this method. I employ the third method which considers a time-series process for the factors and solves the dimensionality problem related to first model. Thus, I estimate the factors using the third method, which is a two-step procedure. First, I estimate the nonparametric factors using the cross-sectional averaging of a principal components method, and second, I estimate the parameters of the state space model as in the first method.

Principal components estimation. In this method, the factors are estimated by cross-sectional averaging of X_t. The idea is that by the weak law of large numbers, the weighted averages of the idiosyncratic disturbances converge to zero and only linear combinations of the factors remain. The principal components estimator is the weighted average estimator of X_t, where the weight is defined as the matrix of scaled eigenvectors of the sample variance matrix of X_t, $\hat{\Sigma}_x$, related to its r largest eigenvalues. The principal components estimator finds the weight matrix which is the solution to the following minimization problem:

$$min_{F_1,\ldots,F_T,\Lambda} V_r(\Lambda, F) = \frac{1}{NT} \sum_{i=1}^{T} (X_t - \Lambda F_t)^{'} (X_t - \Lambda F_t) \tag{1.9}$$

subject to the normalization $N^{-1}\Lambda^{'}\Lambda$.

The key result is that the principal components estimator of the space spanned by the estimated factors in this method is consistent, even when allowing for weak serial and cross-correlation in the idiosyncratic disturbances, and in addition, if the number of observations is sufficiently large, then the estimated factors can be used as data in subsequent regressions.

State space model with static factors. Given the static factors $\hat{F}_t$, estimated by the principal component model, the parameters of the state space model in the "static form" of the DFM, (1.7)-(1.8), are estimated by a regression of X_t on $\hat{F}_t$, and the residuals from this regression are used to estimate the univariate autoregressive dynamic process of idiosyncratic errors. The VAR coefficients $\psi(L)$ are estimated by a regression of $\hat{F}_t$ onto its lags, and the variance of $G\eta_t$ is estimated by the residuals from this VAR.

Determining the number of static factors (r). As mentioned above, the static factors summarize the information contained in the vector of observable variables X_t. Bai and Ng (2002) propose an estimator for the number of static factors (r) based on information criteria, which trades off the benefit of including an additional factor in increasing the explained variation and the cost of increased sampling variability for estimating another factor. They estimate the number of static factors by minimizing:

$$IC(r) = lnV(\hat{\Lambda}, \hat{F}) + rg(N, T) \tag{1.10}$$

where the first term is the least squares objective function (Eq.1.9) evaluated at the principal components estimators $(\hat{\Lambda}, \hat{F})$, and $rg(N, T)$ is a penalty factor. According to Bai and Ng (2002), $g(N, T)$ can be considered as:

$$g(N, T) = (N + T)ln(min(N, T))/(NT) \tag{1.11}$$

which does well in simulations.

Determining the number of dynamic factors (q). The dynamic factors (q) determine the order of autoregressive process for static factors (order of $\Phi(L)$ in Eq.1.8). Amenguel and Watson (2007) propose an estimate of the number of dynamic factors which is based on the observation that the residuals in a regression of X_t on past values of f_t have a factor structure with rank q. Thus, they apply Bai-Ng (2002) information criterion to the sample variance matrix of these residuals to obtain a consistent estimate of the number of dynamic factors.

1.4.3 Specification of Empirical Methodology, Identification, and Estimation

This section lays out the exact empirical methodology for investigating the contribution of fundamentals and speculation to oil price fluctuations. I consider the interaction between oil price fluctuations and a wide range of fundamentals which measure economic activity. Moreover, I consider the interactions among those parts of speculative variables specific to speculators' activity. The empirical methodology which I employ in deriving the common factors is a slightly different version of the methodology applied by Gilchrist et al. (2009). I consider four specifications in this section. First, I explain the baseline specification.

Let X_t, $t = 1, 2, ..., T$, denote a $(n \times 1)$ vector of observations on each variable at time t. The vector of observations (X_t) can be partitioned as $X_t = [X'_{fund,t}, X'_{spec,t}]'$, where $X_{fund,t}$ is the $(n_f \times 1)$ vector whose elements correspond to measures of fundamentals, and $X_{spec,t}$ is the $(n_s \times 1)$ vector whose elements correspond to measures of speculative activity. The information in each partition of the vector of observable variables X_t, can be summarized by a set of latent factors. The $(r_f \times 1)$ vector of "fundamental" factors, $F_{fund,t}$, spans all the information contained in the observed vector $X_{fund,t}$, with $r_f < n_f$. The second group of factors (excess speculation factors), denoted by the $(r_s \times 1)$ vector $F_{exspec,t}$, spans the part of information contained in the observed vector $X_{spec,t}$ which is specifically related to speculative activities, with $r_s < n_s$.

A six-step estimation procedure is used to estimate and identify the fundamental and excess speculation factors, which enables me to investigate the contribution of each partition of observable variables (fundamental and speculative variables) to oil price fluctuations.

First: The $(T \times r_f)$ matrix of factors F_{fund} is estimated as the first r_f principle components of the $(T \times n_f)$ data matrix X_{fund}, while the number of latent static factors (r_f) is chosen according to the Bai and Ng (2002) method. These fundamental factors (F_{fund}) capture the common trend between the first partition of observable variables (fundamentals variables) excluding the idiosyncratic trend which is specific to each of these variables. Hence, these fundamental factors summarize the real activity in the economy.

Second: The number of dynamic factors (q_f) for the fundamental factors (F_{fund}) is estimated using Amenguel and Watson's (2007) procedure. The dynamic process of the fundamental factors then can be described by the following autoregressive process:

$$[F_{fund,t}] = \Phi(L)[F_{fund,t-1}] + [\varepsilon_{f,t}] \tag{1.12}$$

Where $\Phi(L)$ is a matrix polynomial in the lag operator L of finite order q_f, with q_f denoting the number of dynamic fundamental factors, and ε_f is the $(r_f \times 1)$ vector of reduced-form VAR disturbances with a covariance matrix $\Sigma = E[\varepsilon_{ft}\varepsilon^{'}_{ft}]$. The VAR$(q_f)$ model is estimated by OLS using the estimated static factors. Then, I find the estimated r_f factors in F_{fund} at time $t+1$, which represent the expectation of r_f fundamental factors (F_{fund}) at time $t+1$ as of time t.

Third: Each column of the $(T \times n_s)$ data matrix X_{spec} corresponding to the vector of variables in $X_{spec,t}$, (the variables which summarize the speculative activities in the economy) is regressed on the expectation of r_f fundamental factors F_{fund} at time $t+1$ as of time t, $E_t(F_{fund,t+1})$, estimated in the second step in Eq.1.12, with $\hat{E}$ denoting the corresponding $(T \times n_s)$ matrix of OLS residuals. The residuals $(\hat{E})$ are orthogonal to the fundamental common factors, and capture the variation in speculative variables which does not originate from real economic activity. The reason I consider the fundamental factors at time $t+1$ rather than time t, is that efficient markets imply that the speculators' activities are formed based on the expectation of real activity in the next period. Thus the expectation of change in the real activity variables in the next period affects the fluctuation of speculative variables at time t.

Fourth: The $(T \times r_s)$ matrix of factors F_{exspec} is estimated as the first r_s principle components of the data matrix $\hat{E}$ from the second step, and the number of latent static factors (r_s) is chosen according to the Bai and Ng (2002) method. Since the residuals from the third step are orthogonal to fundamental factors (F_{fund}), the excess speculation factors (F_{exspec}) estimated in this step are orthogonal to F_{fund} by construction and represent the common factor of speculators' activities which is not related to real economic activity.

Fifth: Assuming the relationship between oil price fluctuations and the common factors is linear, oil price fluctuations are regressed on the estimated orthogonal factors of fundamental and excess speculation.

$$\Delta Oil_t = \beta_f F_{fund,t} + \beta_s F_{exspec,t} + e_t. \tag{1.13}$$

Sixth: I decompose the variance of oil price fluctuations using the common factors estimated and find the relative contribution of each of these factors to oil price fluctuations.

$$var(\Delta Oil_t) = \beta_f^2 var(F_{fund,t}) + \beta_s^2 var(F_{exspec,t}) + var(e_t). \qquad (1.14)$$

Variables Entering Dynamic Factor Model

The variables included in $X_{fund,t}$ are macroeconomic and financial variables which can be classified into four broad categories. These are economic activity indicators, inflation indicators, risk-free interest rates, and financial indicators. See Appendix A for data sources.

Included as economic activity variables are: (1) the difference of the civilian unemployment rate; (2) the log-difference of nonfarm payroll employment; (3) the log-difference of industrial production index; (4) the difference in capacity utilization index; (5) the Institute for Supply Management (ISM) diffusion index of activity in the manufacturing sector; (6) the log-difference of real personal consumption expenditures (retail control category); (7) the log-difference of housing starts; (8) and the leading economic indicator index.[6]

For variables capturing price dynamics, I consider two inflation indicators: (1) the log-difference of the core CPI; (2) the log- difference of the core PPI.

To capture risk-free interest rates I use: seven interest rates which include the entire term structure of interest rates. In particular, the effective federal funds rate, and

[6]The leading index predicts the six-month growth rate of the coincident index. In addition to the coincident index, the models include other variables that lead the economy: housing permits (1 to 4 units), initial unemployment insurance claims, delivery times from the Institute for Supply Management (ISM) manufacturing survey, and the interest rate spread between the 10-year Treasury bond and the 3-month Treasury bill.

constant maturity Treasury yields at 6 months, 1 year, 2 years, 3 years, 5 years, and 10 years. To obtain the approximate stationarity, I convert the nominal yields into real terms.[7]

Finally, to capture financial market activity I use: (1) the total value-weighted excess market return; (2) the implied volatility on the $S\&P500$ index options (VIX) to capture uncertainty in the equity market; (3) the difference between Moody's Baa and Aaa rated corporate bond yields to capture aggregate default risk in the economy and expected excess returns on stocks and bonds; (4) the implied volatilities on Eurodollar interest rate, a measure of uncertainty of movement in short-term interest rates; (5) the implied volatilities on 10-year Treasury note futures, a measure of uncertainty of movements in long-term interest rates; (6) the log-difference of the trade-weighted exchange value of the dollar against major currencies; and (7) the difference between the 5-year swap rate and the yield on the 5-year Treasury note, a measure of liquidity.

How do I measure speculative activity? As I explained in section 1.3, both spot prices and futures prices react to the same economic forces as economic theory predicts. Thus, the futures risk premium, which is modeled in terms of the sum of the excess market return and the futures excess return, can be employed for explaining the fluctuation in oil spot prices. In this case, the above fundamental variables (X_{fund})

[7]Following Gilchrist et al. (2009), I employ both the realized inflation and survey measures of inflation expectations to convert the nominal yields to real terms. The realized inflation is defined as the difference between the log of the core CPI price index and its value 12 months earlier. The inflation expectation, collected by the Survey of Professional Forecasters (SPF), is at a quarterly frequency and includes the 1-year and 10-years ahead expected CPI inflation. Thus, I obtain the monthly estimates of inflation expectations from a linear interpolation of quarterly inflation expectations. Moreover, the expected inflation for 2 years, 3 years and 5 years is calculated by using the weighted average of 1-year ahead and 10-year ahead expected inflation. Simply, for calculating the 3-year expected inflation, I use the weights of 0.7 and 0.3 respectively on 1-year ahead and 10-year ahead expected inflation. Then, the real federal funds rate is measured as the difference between the nominal rate and realized inflation. The real 6-month Treasury yield is measured as the difference between the nominal yield and the equally weighted average of the realized inflation and the one-year ahead expected inflation. Remaining real Treasury yields, are obtained by the difference between the nominal yields and the appropriate expected inflation rate discussed above.

summarizing the real economy capture both excess market returns and some portion of futures excess returns which respond to fundamentals. The remaining part of the futures excess return is assumed to be driven by speculative activity.

The vector of X_{spec} contains 11 variables which reflect the information about speculators' activity in the economy. They are: (1) the variation in the aggregate level of fundamental hedging demand for oil, which can be proxied for by aggregate commodity producer default risk (EDF); (2) the realized variance of oil futures returns (RV) as a variable for time variation in the quantity of risk in the commodity market; (3) the implied crude oil volatility index (OVX), a measure of quantity of risk in the commodity market; (4) the interaction between expected default frequency of oil companies and the variance of oil futures returns ($EDF \times RV$); (5) the interaction between expected default frequency of oil companies and the implied crude oil volatility index ($EDF \times OVX$); (6) the variation in the aggregate level of fundamental hedging demand, which can be proxied for by the Zmijewski-score (Zm_score) by using the firm-level balance-sheet variables; (7) the interaction between the Zmijewski-score of oil companies and the variance of oil futures returns ($Zm\_score \times RV$); (8) the interaction between the Zmijewski-score of oil companies and the implied crude oil volatility index ($Zm\_score \times OVX$); (9) the risk tolerance of speculators, which can be proxied for by growth in intermediaries' assets relative to household asset growth as a measure of speculators' ease of access to capital in aggregate (γ_s^{-1}); (10) the interaction between expected default frequency of oil companies and the risk tolerance of speculators ($EDF \times \gamma_s^{-1}$); (11) the interaction between the Zmijewski-score of oil companies and the risk tolerance of speculators ($Zm\_score \times \gamma_s^{-1}$).

Thus in the above baseline specification, the vector X_{fund} contains 24 macroeconomic and financial time-series. The 11 elements of vector X_{spec} correspond to the variables that can summarize the speculative activities in the economy. Recalling my procedure, I first find the common factors of X_{fund}, and then extract the part of X_{spec} that is orthogonal to X_{fund}. So it is only speculative activity unrelated to fundamentals that I call "excess speculation".

In addition to the baseline specification, for the robustness check of the results, I consider two other specifications, including,

Specification 1: I incorporate global economic activity by including two proxies for global economic activity in the vector $X_{fund,t}$. Following Juvenal and Petrella (2011), the first proxy which I consider is world industrial production and the second proxy is dry cargo bulk freight rates as proposed by Kilian (2009).

Specification 2: I augment Specification 1 with world oil supply. In fact, I incorporate two proxies for global economic activity as well as world oil supply in the vector $X_{fund,t}$.

1.5 Data

The data are from different sources mentioned in Appendix A, including Federal Reserve Bank of St. Louis, Survey of Professional Forecasters (SPF), the Center for Research in Security Prices ($CRSP$)-Compustat database and the Center for Research in Security Prices ($CRSP$) (in Warton Research Data Services ($WRDS$)), Chicago Board Options Exchange ($CBOE$), U.S. Energy Information Administration (EIA), the U.S. Flow of Funds data, Economagic website, and Lutz Kilian's website [8]. It

[8]http://www.economagic.com
http://www-personal.umich.edu/ lkilian/paperlinks.html

includes data over the period 1986-2013 for most of the series, while the data for some series is available only for a subset of the period. In addition, I construct proxies for some of the variables described in Appendix B. The data and proxies for variables have been tested for stationarity and those which are integrated of order one ($I(1)$) have been transformed to difference out unit roots and trends. Then, as is standard in using principal component models, the data have been standardized to have mean zero and unit standard deviation.

1.6 Empirical Results

The empirical results are shown in Table 1.1-1.5. It includes the results for different specification discussed in section 1.4.3 and for different sample periods based on the data availability. As mentioned in the Introduction, the variance of the oil price has increased substantially after 2003. To study how oil price fluctuations have evolved over time in response to changes in fundamental and speculative activities, I also consider two distinct periods, the low volatility period in the oil price (before 2003, subsample 1) and the high volatility period in the oil price (after 2003, subsample 2). This can then show a possible structural break in the contribution of fundamentals and speculative activities to oil price fluctuations.

Table 1.1 shows the estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Baseline specification, including 19 fundamental variables available for the whole sample period 1986-2013. The contribution of fundamentals is about 3% for the whole period, 1.1% for the subsample 1, and 8.3% for the subsample 2. The contribution of second lagged of excess speculation factors to oil price fluctuations is about 2% for the whole sample, 1.1% for the subsample

1, and 27.5% for the subsample 2. The results show an increase in contribution of fundamental factors and a surge in contribution of excess speculation factors to oil price fluctuations after 2003.

Table 1.2 shows the estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Specification 1, augmenting two proxies of global economic activity to fundamental variables in Baseline specification. The sample period changes to 1994-2013 due to availability of data on world industrial production. The contribution of fundamentals is about 12.9% for the whole period, 3.6% for the subsample 1, and 9.8% for the subsample 2. The contribution of second lagged of excess speculation factors to oil price fluctuations is about 5.1% for the whole sample, 4.5% for the subsample 1, and 16.2% for the subsample 2. The results show an increase in contribution of fundamental factors and excess speculation factors to oil price fluctuations after 2003.

Table 1.3 shows the estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Specification 2, augmenting world oil supply to fundamental variables in Specification 1. The contribution of fundamentals is about 12.6% for the whole period, 2.1% for the subsample 1 (1994-2003), and 11.2% for the subsample 2 (2004-2013). The contribution of second lagged of excess speculation factors to oil price fluctuations is about 5.1% for the whole sample, 4.3% for the subsample 1, and 18.1% for the subsample 2. The results show an increase in contribution of fundamental factors and excess speculation factors to oil price fluctuations after 2003.

Table 1.4 shows the estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Baseline specification, including 22 fundamental variables. The sample starts from 2004:10 since the data on three of the fundamental variables including Real personal consumption expenditures, VIX, and a measure of liquidity (defined as the difference between the 5-year swap rate and the yield on the 5-year Treasury note) is available after 2004:10. The contribution of fundamentals and first lagged of excess speculation factors to oil price fluctuations is about 19.8% and 21.2% respectively.

Table 1.5 shows the estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Baseline specification, including all 24 fundamental variables and 11 speculative variables mentioned in the empirical methodology section. The data on all fundamental variables and speculative variables, mentioned in the empirical methodology section, are available after 2008:1. The contribution of fundamentals is about 3.7% and the contribution of contemporaneous, first lagged , and second lagged excess speculation factors to oil price fluctuations is about 46.2%, 25.2%, and 18.1% respectively.

The results in all tables lay out a structural break in the contribution of fundamentals and speculative activities to oil price fluctuations after 2003, and an increased contribution of speculators in the oil market after 2003 compared to before 2003.

1.7 Conclusion

The results in all of the specifications considered above lay out evidence of "excess speculation" in the oil market after 2003, which solely originates from speculators' activity, and it is not a function of fundamentals. The results show excess speculation is the major driver of oil price fluctuations after 2003.

The potential channel through which speculators can impact the oil market and the dynamics of futures and spot oil prices is through the hedging demand of producers. Oil producers are risk-averse and speculators face capital constraints. When the capital constraints bind, this limits the ability of oil producers to hedge their production against price risk, which exposes them to more price risk. This leads to greater volatility in oil price fluctuations. Thus, in conclusion, if speculators buy lots of futures, and capital constraints become binding, this might lead to greater volatility in oil price fluctuations. This, in turn, increases the asset volatility of oil producers, which leads to an increase in the expected default frequencies of oil producers exposing them to even more risk and default in the subsequent periods. Hence, the effect can even be explosive.

Moreover, there might be a case that the speculators' activity creates further volatilities by rolling their entire position forward. In this case, the limits on the ability of oil producers to hedge their production have larger impacts on oil price fluctuations as a result of greater volatility induced by speculators.

The channel described above can be considered as excess speculation if unrelated to fundamentals, and the results in this paper show evidence of that. The results imply that this market should be regulated in order to restrict the influence of speculators' activity on oil price dynamics, helping the oil market to function properly.

Bibliography

[1] Acharya, V.V., Lochstoer, L.A., Ramadorai, T., 2013. "Limits to Arbitrage and Hedging: Evidence from Commodity Markets," *Journal of Financial Economics*,109: 441-465.

[2] Alquist, R., Kilian, L., 2010. "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics*, 25: 539-573.

[3] Amenguel, D., Watson, M.W., 2007. "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel," *Journal of Business and Economic Statistics*,91-96.

[4] Ang, A., Bekaert, G., 2007. "Stock Return Predictability: Is It True?" *Review of Financial Studies*, 20: 651-707.

[5] Bai, J., Ng, S., 2002. "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70: 191-221.

[6] Bernanke, B. S., 2004. "Oil and the Economy," *Speech, The Federal Reserve Board*, October 2004.

[7] Black, F., Scholes, M., 1973. "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81: 637-659.

[8] De Roon, F. A., Nijman, T. E., Veld, C., 2000. "Hedging Pressure Effects in Futures Markets?" *Journal of Finance*, 55: 1437-1456.

[9] Engle, R. F., 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50: 987-1007.

[10] Etula, E., 2009. "Broker-dealer Risk Appetite and Commodity Returns," *Staff Reports 406, Federal Reserve Bank of New York*.

[11] Fama, E. F, French, K. R., 1989. "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25: 23-49.

[12] Fattouh, B., Kilian, L., Mahadeva, L., 2012. "The Role of Speculation in Oil Markets: What Have We Learned So Far?" Working paper. Federal Reserve Bank of St. Louis.

[13] Gilchrist, S., Yankov, V., Zakrajsek, E., 2009. "Credit Market Shocks and Economic Fuctuations: Evidence from Corporate Bond and Stock Markets," *Journal of Monetary Economics*, 56: 471-493.

[14] Hamilton, J.D., 1994. "Time Series Analysis," *Princeton University Press*, 1994.

[15] Hamilton, J.D., 2009b. "Understanding Crude Oil Prices," *Energy Journal*, 30: 179-206.

[16] Haushalter, G.D., 2000. "Financing Policy Basis, Risk, and Corporate Hedging: Evidence from Oil and Gas Producers," *Journal of Finance*, 55: 107-152.

[17] Hirshleifer, D., 1988. "Residual Risk, Trading Costs, and Commodity Risk Premia?" *Review of Financial Studies*, 1: 173-193.

[18] Hirshleifer, D., 1990. "Hedging Pressure and Futures Price Movements in a General Equilibrium Model," *Econometrica*, 58: 411-428.

[19] Juvenal, L., Petrella, L., 2011. "Speculation in the Oil Market," Working paper. Federal Reserve Bank of St. Louis.

[20] Kilian, L., 2009. "Not all Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 99(3): 1053-1069.

[21] Kilian, L., Murphy, D.P., 2011. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," Working paper. University of Michigan.

[22] Kilian, L., Murphy, D.P., 2014. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," *Journal of Applied Econometrics*, 29: 454-478.

[23] Lombardi, M.J., Van Robays, I., 2011. "Do Financial Investors Destabilize the Oil Price," Working paper. European Central Bank.

[24] Masters, M.W. 2008. "Testimony before the U.S. Senate Committee on Homeland Security and Governmental Affairs," May 2008.

[25] Merton, R.C., 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29: 449-470.

[26] Stock, J.H., Watson M.W., 2010. "Dynamic Factor Models," *The National Bureau of Economic Research, 2010.*

[27] Sun, Z., Munves, D., Hamilton, D., 2012. "Public Firm Expected Default Frequency EDF^{TM} Credit Measures: Methodology, Performance, and Model Extensions," *Capital Markets Research, Moody's Analytics*, June 2012.

[28] Zmijewski, M.E., 1984. "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," *Journal of Accounting Research*, 22: 59-82.

Table 1.1: Baseline specification. Estimated contribution of fundamentals and excess speculation factors to oil price fluctuations, including 19 fundamental variables available for the whole sample.

Dependent variable: log spot oil price changes	1986-2013	1986-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	2.99%	1.10%	8.28%
1-month ahead forecasted fundamental factors	3.32%	1.10%	9.26%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	4.05%	0%	2.27%
First lagged excess speculation factors	2.03%	0.33%	14.12%
Second lagged excess speculation factors	1.99%	1.08%	27.46%
Number of static fundamental factors (r_f)	2	2	2
Number of dynamic fundamental factors (q_f)	2	1	4
Number of static excess speculation factors (r_s)	3	1	1

Notes: Contributions are the share of variance of oil price fluctuations accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(\Delta Oil_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(\Delta Oil_t)}$*). The estimation includes 19 fundamental variables available for the whole sample.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund})*.* r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR(*q_f*) model for the estimated* r_f *static fundamental factors* F_{fund}*.*

Table 1.2: Specification 1. Estimated contribution of fundamentals and excess speculation factors to oil price fluctuations, augmenting two proxies of global economic activity to fundamental variables in Baseline specification.

Dependent variable: log spot oil price changes	1994-2013	1994-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	12.92%	3.65%	9.84%
1-month ahead forecasted fundamental factors	12.59%	8.49%	12.44%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	5.29%	0.10%	19.19%
First lagged excess speculation factors	3.69%	0.69%	13.21%
Second lagged excess speculation factors	5.11%	4.47%	16.17%
Number of static fundamental factors (r_f)	4	4	2
Number of dynamic fundamental factors (q_f)	4	5	5
Number of static excess speculation factors (r_s)	3	1	3

Notes: Contributions are the share of variance of oil price fluctuations accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(\Delta Oil_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(\Delta Oil_t)}$*). The sample starts from* $1994:10$ *due to availability of data on world industrial production.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund}). r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR*(q_f) *model for the estimated* r_f *static fundamental factors* F_{fund}.

Table 1.3: Specification 2. Estimated contribution of fundamentals and excess speculation factors to oil price fluctuations, augmenting world oil supply to fundamental variables in Specification 1.

Dependent variable: log spot oil price changes	1994-2013	1994-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	12.56%	2.08%	11.21%
1-month ahead forecasted fundamental factors	13.57%	4.73%	18.20%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	5.09%	0.02%	16.36%
First lagged excess speculation factors	4.10%	0.53%	12.62%
Second lagged excess speculation factors	5.09%	4.29%	18.10%
Number of static fundamental factors (r_f)	4	3	3
Number of dynamic fundamental factors (q_f)	4	5	4
Number of static excess speculation factors (r_s)	3	1	3

Notes: Contributions are the share of variance of oil price fluctuations accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(\Delta Oil_t)}$ and $\frac{\beta_s^2 var(F_{exspec})}{var(\Delta Oil_t)}$). The sample starts from 1994 : 10 due to availability of data on world industrial production. r_f shows the number of "fundamental" factors (F_{fund}) which span all the information contained in the observed fundamental variables (X_{fund}). r_s represents the number of excess speculation factors (F_{exspec}) which span the part of information contained in the observed speculative variables (X_{spec}) unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors (q_f) shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of r_f fundamental factors (F_{fund}) at time $t+1$ as of time t is found by estimating the VAR(q_f) model for the estimated r_f static fundamental factors F_{fund}.

Table 1.4: Estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Baseline specification, including 22 fundamental variables.

Dependent variable: log spot oil price changes	2004-2013
Contribution of fundamentals (F_{fund})	
Contemporaneous fundamental factors	19.80%
1-month ahead forecasted fundamental factors	19.84%
Contribution of excess speculation factors (F_{exspec})	
Contemporaneous excess speculation factors	19.52%
First lagged excess speculation factors	21.17%
Second lagged excess speculation factors	11.93%
Number of static fundamental factors (r_f)	3
Number of dynamic fundamental factors (q_f)	5
Number of static excess speculation factors (r_s)	3

Notes: Contributions are the share of variance of oil price fluctuations accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(\Delta Oil_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(\Delta Oil_t)}$*). The data on all three fundamental variables including Real personal consumption expenditures,* VIX*, and a measure of liquidity (defined as the difference between the 5-year swap rate and the yield on the 5-year Treasury note) are available after* 2004 : 10*. The results are based on Baseline specification including 22 fundamental variables.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund})*.* r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *determines the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR(*q_f*) model for the estimated* r_f *static fundamental factors* F_{fund}*.*

Table 1.5: Estimated contribution of fundamentals and excess speculation factors to oil price fluctuations in Baseline specification, including all 24 fundamental variables and 11 speculative variables mentioned in the empirical methodology section.

Dependent variable: log spot oil price changes	2008-2013
Contribution of fundamentals (F_{fund})	
Contemporaneous fundamental factors	3.75%
1-month ahead forecasted fundamental factors	3.75%
Contribution of excess speculation factors (F_{exspec})	
Contemporaneous excess speculation factors	46.20%
First lagged excess speculation factors	25.10%
Second lagged excess speculation factors	18.12%
Number of static fundamental factors (r_f)	2
Number of dynamic fundamental factors (q_f)	1
Number of static excess speculation factors (r_s)	4

Notes: Contributions are the share of variance of oil price fluctuations accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(\Delta Oil_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(\Delta Oil_t)}$*). The data on all 24 fundamental variables and 11 speculative variables, mentioned in the empirical methodology section, are available after* 2008 : 1. *The results are based on Baseline specification including 24 fundamental variables and 11 speculative variables.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund}). r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *determines the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR*(q_f) *model for the estimated* r_f *static fundamental factors* F_{fund}.

Appendices

Appendix A

Data sources

Table A.1: Data sources

Variables	Source	Start Date	End Date
Economic activity			
Civilian unemployment rate	Federal Reserve Bank of St. Louis	1986	2013
Nonfarm payroll employment	Federal Reserve Bank of St. Louis	1986	2013
Industrial production index	Federal Reserve Bank of St. Louis	1986	2013
Capacity utilization index	Federal Reserve Bank of St. Louis	1986	2013
ISM diffusion index of activity in the manufacturing sector	Federal Reserve Bank of St. Louis	1986	2013
Real personal consumption expenditures	Federal Reserve Bank of St. Louis	1999	2013
Housing starts	Federal Reserve Bank of St. Louis	1986	2013
The Leading economic indicator index	Federal Reserve Bank of St. Louis	1986	2013
Inflation indicator			
Core CPI	Federal Reserve Bank of St. Louis	1986	2013
Core PPI	Federal Reserve Bank of St. Louis	1986	2013
Risk-free interest rates			
Effective federal funds rate	Federal Reserve Bank of St. Louis	1986	2013
6-month maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
1-year maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
2-year maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
3-year maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
5-year maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
10-year maturity Treasury yields	Federal Reserve Bank of St. Louis	1986	2013
Inflation expectation	Survey of Professional Forecasters (SPF)	1986	2013
Financial market activity			
Total value-weighted excess market return	Warton Research Data Services ($WRDS$)	1986	2013

Variables	Source	Start Date	End Date
Implied volatility on the S&P500 index options	Chicago Board Options Exchange ($CBOE$)	1990	2013
Moody's Aaa rated corporate bond yields	Federal Reserve Bank of St. Louis	1986	2013
Moody's Baa rated corporate bond yields	Federal Reserve Bank of St. Louis	1986	2013
Implied volatilities on Eurodollar interest rate	Chicago Board Options Exchange ($CBOE$)	2007	2013
Implied volatilities on 10-year Treasury note futures	Chicago Board Options Exchange ($CBOE$)	2008	2013
Trade-weighted exchange value of the dollar	Federal Reserve Bank of St. Louis	1986	2013
5-year swap rate	Federal Reserve Bank of St. Louis	2004	2013
Yield on the 5-year Treasury note	Federal Reserve Bank of St. Louis	1986	2013
Global economic activity			
World industrial production	IMF	1994	2013
dry cargo bulk freight rates	Lutz Kilian's website	1986	2013
World oil production	Economagic website	1986	2013
Speculation activity			
Oil producers' default risk	Warton Research Data Services ($WRDS$)	1986	2013
Oil futures price	U.S. Energy Information Administration (EIA)	1986	2013
Implied crude oil volatility index			
Oil producers' Zmijewski-score	Warton Research Data Services ($WRDS$)	1986	2013
Risk tolerance of speculators	Federal Reserve Bank of St. Louis	1986	2013
Oil spot price	Federal Reserve Bank of St. Louis	1986	2013

Notes: Except the implied crude oil volatility index, I estimate all of the speculative variables using the data available on the sources mentioned above.

Appendix B

Proxies for Variables

Proxying for Fundamental Hedging Demand

As discussed in section 1.3, one of the components which affects the futures return in commodity markets besides the excess return on equities is the producers' risk-aversion which drives producers to hedge their production. Acharya et al. (2013), analyzing 2400 firm quarterly reports, show that 69.8% of oil companies hedge their production by going short and 1% of these companies initiate long positions.

Haushalter (2000) study the hedging policies of oil and gas producers over the 1992-1994 period. He shows the extent of hedging is positively associated with financial leverage, measured as the ratio of total debt to total asset. This indicates risk management in oil companies to reduce the likelihood of financial distress. Moreover, Acharya et al. (2013) show that the degree of oil derivative hedging is positively related to default risk measures, the Zmijewski (1984) score and expected default frequency (EDF). The likelihood that a firm defaults can be a guide about the firm's hedging demand and as a result, oil hedging demand can be proxied for by the default risk of oil companies.

Following Acharya et al. (2013), I employ both the Zmijewski (1984) score and EDF as proxies for fundamental hedging demand of oil producers. Both fundamental credit risk (Zmijewski-score), and structural credit risk (EDF) are computed based on the analysis of a firm's balance sheets. The main difference between these two is that Zmijewski-score is an accounting value-based, while EDF is a forward-looking measure of risk based on the market value.

I estimate the Zmijewski-score and EDF for oil and gas producer firms with SIC code 1311 (490 firms) using the balance-sheet data from the Center for Research in Security Prices ($CRSP$)-Compustat database. Then I aggregate the firm's Zmijewski-score and EDF estimated to obtain an aggregate indicator of oil producers' expected default, and as a result, an aggregate indicators of overall producers' hedging demand.

I find the Zmijewski score employing the firm-level balance-sheet variables. Each firm's Zmijewski-score is calculated as follows.

$$\begin{gathered} Zmijewski - score = -4.3 - 4.5^{*} NetIncome/TotalAssets \\ +5.7^{*} TotalDebt/TotalAssets \\ -0.004^{*} CurrentAssets/CurrentLiabilities \end{gathered} \tag{B.1}$$

The market-based measure of credit risk, which I employ, is the expected default frequency (EDF). I estimate the EDF for each firm applying the empirical method used by Moody's KMV.

Public firm expected default frequency EDF is one of the forward-looking probabilities of default estimates for publicly traded companies and is a modification of Merton's approach for structural credit risk models considering more realistic assumptions. The EDF is a cardinal measure of credit risk, rather than a rank ordering of credit risk.

It is motivated by the conceptual underpinning that default is highly likely to occur when the market value of a firm's assets is insufficient to cover the book value of its liabilities at some future date. In other words, the model computes the probability of default as the probability that a firm's assets will fall below its liabilities at a specific time horizon. Thus, structural credit risk (EDF), like fundamental credit risk (Zmijewski-score), is computed based on the analysis of a firm's balance sheets. The advantage, however, is that the EDF, by utilizing financial market information on firm's value, provides an up-to-date measure of firm's credit risk.

There are several difficulties in estimating the EDF, including estimating the market value of firm's asset and estimating the volatility of the firm's assets.

Following the empirical strategy of Moody's KMV, I employ a four-step procedure for estimating public firm EDF: (1) utilizing an iterative procedure to estimate the asset value and the empirical volatility of firm's asset; (2) computing the default point for a one-year time horizon; (3) computing the firm's distance to default (DD) using the estimates from steps 1 and 2; (4) finding EDF from DD in step 3, based on assumption of the normally distributed of DDs.

In Appendix C, in more details, I explain the theoretical underpinnings of the EDF, the difficulties in estimating the expected default frequency (EDF) and the exact feature of Moody's KMV method for estimating EDF.

Estimates for Volatility of Oil Futures Returns

I first define the measure of oil futures returns. Following the methodology of Acharya et al. (2013), I construct oil futures returns at the end of month t as the one-period price percentage change in the nearest-to-maturity contract that would not expire during the month $t+1$. Thus, the oil futures return is calculated as:

$$\frac{F_{t+1,T} - F_{t,T}}{F_{t,T}}. \tag{B.2}$$

where $F_{t,T}$ is the oil futures price at the end of month t on the nearest contract that expires after the end of month $t+1$, at time T, and $F_{t+1,T}$ is the price of the same contract at the end of month $t+1$.

Then, I estimate the conditional variance (the forecasted or realized variance) at time t of oil futures return from time t to $t+1$ using the autoregressive conditional heteroskedastic ($ARCH$) model introduced by Engle (1982). The conditional variance of oil futures returns evolves according to the autoregressive process given by:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \alpha_3 \hat{\varepsilon}_{t-3}^2 + \nu_t. \tag{B.3}$$

Table B.1: The realized variance at time t of oil futures returns.

Regressor	$\hat{\varepsilon}_{t-1}^2$	$\hat{\varepsilon}_{t-2}^2$	$\hat{\varepsilon}_{t-3}^2$
Dependent variable: variance at time t of oil futures returns ($\hat{\varepsilon}_t^2$)	0.1116	0.0701	0.2048

where $\hat{\varepsilon}_t$ is the error in forecasting (modeling) oil futures returns as an AR process. As shown in table B.1, all of the coefficients on the above unrestricted AR process using squares of the estimated residuals are positive and ν_t is bounded from below by $-\alpha_0$, insuring that the conditional variance is never negative. Moreover, the sum of coefficients are less than 1, ensuring that the process is covariance-stationary and stable.

Proxy for Speculator Risk Tolerance

Following Etula (2009) and Acharya et al. (2013), I consider a growth in intermediaries' (aggregate broker-dealer) assets relative to household asset growth as a proxy for the speculator risk tolerance (speculator' risk-aversion) in aggregate.

Appendix C

The Challenges and the Detailed Method for Estimating Firm Expected Default Frequency (EDF)

There are some difficulties for estimating the expected default frequency (EDF), including.

(1) Estimating the market value of firm's asset: A company's assets worth in the balance sheet reflects the book value cost of the asset at the time of purchase less accumulated depreciation. The book value of the asset does not show an up-to-date value of the asset, which reflects the change in an asset's potential to generate cash in the future. Since markets are forward looking, they can provide a better alternative for estimating the value of a firm, but a company's assets do not trade on the market. In this case, a company's equity value, which trades on the market, can be employed for estimating the unobservable market value of a company's asset.

(2) Estimating the volatility of the firm's assets: The default risk is higher for the businesses which firm's assets are more volatile. In this case, there is a higher probability that the value of the assets fall below the book value of liabilities. The asset volatility is not measurable since the market value of a firm's asset is unobservable.

The theory behind the EDF model for estimating market value and volatility of asset is based on Black-Scholes-Merton structural model (Black and Scholes, 1973, and Merton, 1974). It considers the following stochastic process for the dynamic of market value of a firm's assets

$$dA = \mu A dt + \sigma_A A dW \tag{C.1}$$

where μ is the expected growth rate of the firm's asset value, σ_A is the asset volatility, and W represent a standard Brownian motion.[1] By the geometric Brownian motion assumption, the log asset value is normally distributed

$$lnA_1 = N\left(lnA_0 + (\frac{\mu - \sigma_A^2}{2}), \sigma_A^2\right) \tag{C.2}$$

The likelihood that a firm's asset value (A) fall below its default point (X, which is computed based on the firm's book value of liabilities) in the specific period,

[1]The standard Brownian motion is a continuous-time stochastic process of a random walk; Random walk is a discrete process which is only defined at integer values of t, while the standard Brownian motion is defined at a finer and finer grid of dates (continuous time) at which, the value of this process, $W(t)$, is continuous with probability of 1.

$Pr(lnA_1 < lnX)$ is

$$Probabilityofdefault = \Phi \left[-\frac{ln(\frac{A}{X}) + (\mu - \frac{\sigma_A^2}{2})}{\sigma_A} \right] \tag{C.3}$$

or by using the definition for the distance to default (DD) :

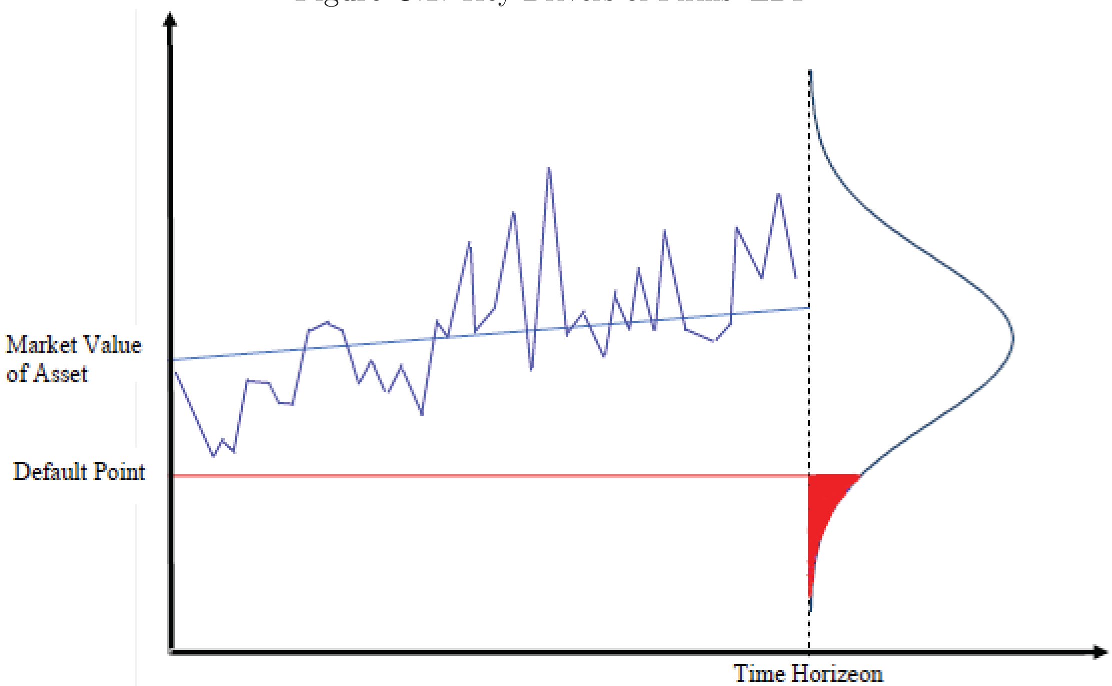
$$Probabilityofdefault = \Phi[-DD] \tag{C.4}$$

For simplicity, by dropping the second term of the numerator above, the DD can be written as:

$$DD = \left[-\frac{ln(A) - ln(X)}{\sigma_A} \right] \tag{C.5}$$

The distance to default or a firm incorporates three important pieces of information: (1) the market value of assets; (2) the default point; and (3) the asset volatility. The distance to default or a firm is the difference between expected market value of the asset and the default point at specific horizon date in the future, standardized by its asset volatility (its business risk); it is a single rank ordering static reflecting the probability of a firm to default (shown by the black area below the default point in Figure C.1).

Figure C.1: Key Drivers of Firms' EDF



A company's unobservable asset value and asset volatility can be derived from its equity market returns. The equity of a company can be considered the same as a call option on the company's assets. In this case, the Black-Scholes option pricing formula for a call option relates the observable equity value to the unobservable asset value and asset volatility for a firm which is financed by equity and a zero-coupon bond (X). Considering E as the value of equity, as the expected growth rate of the firm's asset value, and T as the time horizon, the Black-Scholes formula for the value of equity is

$$E = A_0\Phi(d_1) - e^{-rT}X\Phi(d_2) \tag{C.6}$$

where

$$d_1 = \frac{ln(\frac{A_0}{X}) + (r + \frac{\sigma_A^2}{2})T}{\sigma_A\sqrt{T}} \tag{C.7}$$

$$d_2 = d_1 - \sigma_A\sqrt{T} \tag{C.8}$$

This formula relates the value of a firm's equity (E) to its asset value (directly) and its asset volatility (indirectly, through the normal distribution). Thus, the value of the option, equity value, can be employed to infer the implied firm asset value and volatility. Then, the estimates of firm asset value and volatility can be used for computing the firm's distance to default formula (B.5), and probabilities of default formula (B.3).

Following the empirical strategy of Moody's KMV, I employ a four-step procedure for estimating Public firm expected default frequency (EDF):

First: I utilize an iterative procedure to estimate the asset value and the empirical volatility of firm's asset. For an initial guess of asset volatility, I find a time-series of asset values using the asset value equation. Then, I find the asset volatility of the asset values generated. I consider this new asset volatility for finding a new time-series of asset values. I iterate this process until the difference between asset volatilities estimated in two iterations is so small, or in other words, until the asset volatility estimated converges. The time-series of asset values in the final iteration consider as the firm's asset values.

Second: I compute the default point for a one-year time horizon, as 100% of short-term liabilities plus one-half of long-term liabilities for non-financial firms.

Third: I compute the firm's distance to default (DD) according to Eq.C.5, using the estimates of asset value and volatility, and the default point from steps 1 and 2.

Fourth: I find EDF from DD in step 3, based on assumption of the normally distributed of DDs, as a result of the geometric Brownian motion assumption for the dynamic of asset values (Eq.C.1).

Chapter 2

Did Speculation Cause Variation in Oil Risk Premia?

2.1 Introduction

The literature lacks a general model of oil futures risk premia that incorporates the notion of time-varying risk premia along with other explanations of oil price fluctuations. Indeed, it is not clear in the literature how to construct such a model.

Moreover, as discussed in section 1.2.1 of chapter 1, one of the six strands in the literature on the role of speculation in the oil market, considers models in which the risk premium in the commodity market evolves over time, and takes it as evidence of speculative activities in oil futures market. Although time-varying risk premia might be suggestive of the influence of financial investors on oil futures markets, there is no evidence of excess speculation based on these models.

It is critical to distinguish between two types of speculative activity, speculation and "excess speculation" as mentioned in chapter 1. Speculation is closely related to fundamentals, provides liquidity to oil markets, and helps the functions of price discovery and risk transfer between the future and the present, while excess speculation is not linked to changes in fundamentals, may prevent the oil market from functioning properly generating changes in the dynamic of oil futures and spot prices unrelated to fundamentals, and is only favorable from a private perspective.

This paper uses a new methodology to incorporate all possible explanatory variables of oil risk premia, and then employs it to examine the contribution of fundamentals and excess speculation to oil risk premia and to analyze if the contribution of fundamentals and excess speculation to oil risk premia has evolved over the period 1986-2013.

In practice, I use a dynamic factor model (DFM) to capture the common component of fluctuations in real activity in the economy ("fundamentals") and overall speculative activities by both pure speculators and oil companies. I then take the residual from a regression of the speculation factors on the expectation of the fundamentals factors. The residuals show the part of overall speculative activity which cannot be explained by expectations of fundamentals, and are orthogonal to fundamental factor by construction. I label the residuals "excess speculation" factors. Finally, I calculate how much of the variation in oil futures risk premia is caused by each groups of these orthogonal factors.

My paper contributes to the literature in two dimensions: First and more importantly, as a novel feature in the literature, I construct a model of oil risk premia which addresses the existence of excess speculation. The methodology introduced in this paper can be employed as a correct way for providing the evidence of an increased speculative activities in recent years based on time-varying risk premia models. Second, I augment other explanatory variables of oil price fluctuations to the model of oil risk premia

The results in this paper lay out evidence of "excess speculation" after 2003 based on time-varying risk premia strand of the literature. As mentioned, this type of speculation originates from pure speculators' activity especially after 2003, and it is not a function of fundamentals.

The reminder of the paper is as follows. Section 2.2 provides an underpinning theory about how I define and classify the variables into two groups of fundamentals and speculation. Section 2.3 discusses the empirical methodology of the paper in detail. Section 2.4 discusses the data sources. Section 2.5 lays out the results, and finally, Section 2.6 concludes.

2.2 The Theoretical Underpinning for the Variables in the Empirical Methodology

For my method, deciding what variables to include in the dynamic factor model (DFM) and classifying them as fundamental and speculation variables are crucial. The futures risk premium reflects the compensation for the uncertainty perceived in the oil market. Therefore, changes in uncertainty in the oil market provide information on changes in oil risk premia. Variables which can gauge the extent of uncertainty in the oil market provide me with the pool of potential explanatory variables for oil risk premia. I then classify these variables into two groups, fundamental and speculative variables.

Two sources of uncertainty in the oil market can be considered. One source of uncertainty is in the situation of the whole economy, and this is measured by the business cycle variables and is reflected in the excess market return. The business cycle variables thus can be considered as fundamental variables. The second source of uncertainty is specific to the oil market, and can be measured by uncertainty in the demand and supply of oil. Variables measuring this second source of uncertainty, capture the overall activity of both groups of speculators, oil producers and pure speculators. Hence, the speculative variables can be considered as variables that distort oil futures risk premia from the risk-adjustment of equity holders and capture the impact and interaction of two groups of speculators (pure speculators and oil producers) on oil futures risk premia. Section 1.3 of chapter 1 discusses the theoretical underpinning for these inclusion and classification in detail.

2.3 Empirical Methodology

The empirical methodology discussed in section 1.4.3 of chapter 1, can be applied for modeling time-varying risk premia along with other explanations of oil price fluctuations, and can address the existence of excess speculation based on these models. Thus, below, I explain the empirical methodology for modeling risk premia.

Let X_t, $t = 1, 2, ..., T$, denote a $(n \times 1)$ vector of observations on each variable at time t. The vector of observations (X_t) can be partitioned as $X_t = [X'_{fund,t}, X'_{spec,t}]'$, where $X_{fund,t}$ is the $(n_f \times 1)$ vector whose elements correspond to measures of fundamentals, and $X_{spec,t}$ is the $(n_s \times 1)$ vector whose elements correspond to measures of speculative activity. The information in each partition of the vector of observable variables X_t, can be summarized by a set of latent factors. The $(r_f \times 1)$ vector of "fundamental" factors, $F_{fund,t}$, spans all the information contained in the observed vector $X_{fund,t}$, with $r_f < n_f$. The second group of factors (excess speculative factors), denoted by the $(r_s \times 1)$ vector $F_{spec,t}$, spans the part of information contained in the observed vector $X_{spec,t}$ which is specifically related to speculative activities, with $r_s < n_s$.

A six-step estimation procedure is used to estimate and identify the fundamental and excess speculation factors, which enables me to investigate the contribution of each partition of observable variables (fundamental and speculative variables) to oil futures return.

First: The $(T \times r_f)$ matrix of factors F_{fund} is estimated as the rst r_f principle components of the $(T \times n_f)$ data matrix X_{fund}, while the number of latent static factors (r_f) is chosen according to the Bai and Ng (2002) method. These fundamental factors (F_{fund}) capture the common trend between the first partition of observable variables (fundamentals variables) excluding the idiosyncratic trend which is specific to each of these variables. Hence, these fundamental factors summarize the real activity in the economy.

Second: The number of dynamic factors (q_f) for the fundamental factors (F_{fund}) is estimated using Amenguel and Watson's (2007) procedure. The dynamic process of the fundamental factors then can be described by the following autoregressive process:

$$[F_{fund,t}] = \Phi(L)[F_{fund,t-1}] + [\varepsilon_{f,t}] \tag{2.1}$$

Where $\Phi(L)$ is a matrix polynomial in the lag operator L of finite order q_f, with q_f denoting the number of dynamic fundamental factors, and ε_f is the $(r_f \times 1)$ vector of reduced-form VAR disturbances with a covariance matrix $\Sigma = E[\varepsilon_{ft}\varepsilon_{ft}^{'}]$. The VAR($q_f$) model is estimated by OLS using the estimated static factors. Then, I find the estimated r_f factors in F_{fund} at time $t+1$, which represent the expectation of r_f fundamental factors (F_{fund}) at time $t+1$ as of time t.

Third: Each column of the $(T \times n_s)$ data matrix X_{spec} corresponding to the vector of variables in $X_{spec,t}$, (the variables which summarize the speculative activities in the economy) is regressed on the expectation of r_f fundamental factors F_{fund} at time $t+1$ as of time t, $E_t(F_{fund,t+1})$, estimated in the second step in Eq.16, with $\hat{E}$ denoting the corresponding $(T \times n_s)$ matrix of OLS residuals. The residuals ($\hat{E}$) are orthogonal to the fundamental common factors, and capture the variation in speculative variables which does not originate from real economic activity. The reason I consider the fundamental factors at time $t+1$ rather than time t, is that efficient markets imply that the speculators' activities are formed based on the expectation of real activity in the next period. Thus the expectation of change in the real activity variables in the next period affects the fluctuation of speculative variables at time t.

Fourth: The $(T \times r_s)$ matrix of factors F_{exspec} is estimated as the first r_s principle components of the data matrix $\hat{E}$ from the second step, and the number of latent static factors (r_s) is chosen according to the Bai and Ng (2002) method. Since the residuals from the third step are orthogonal to fundamental factors (F_{fund}), the excess speculation factors (F_{exspec}) estimated in this step are orthogonal to F_{fund} by construction and represent the common factor of speculators' activities which is not related to real economic activity.

Fifth: Assuming the relationship between oil futures return and the common factors is linear, oil futures returns are regressed on the estimated orthogonal factors of fundamental and excess speculation.

$$OilFuturesReturn_t = \beta_f F_{fund,t} + \beta_s F_{exspec,t} + e_t \tag{2.2}$$

Sixth: I decompose the variance of oil futures returns using the common factors estimated and find the relative contribution of each of these factors to oil futures returns.

$$var(OilFuturesReturn_t) = \beta_f^2 var(F_{fund,t}) + \beta_s^2 var(F_{exspec,t}) + var(e_t) \tag{2.3}$$

Variables Entering Dynamic Factor Model

The variables included in $X_{fund,t}$ are macroeconomic and financial variables which can be classified into four broad categories. These are economic activity indicators, inflation indicators, risk-free interest rates, and financial indicators.

Included as economic activity variables are: (1) the difference of the civilian unemployment rate; (2) the log-difference of nonfarm payroll employment; (3) the log-difference of industrial production index; (4) the difference in capacity utilization index; (5) the Institute for Supply Management (ISM) diffusion index of activity in the manufacturing sector; (6) the log-difference of real personal consumption expenditures (retail control category); (7) the log-difference of housing starts; (8) and the leading economic indicator index.[1]

For variables capturing price dynamics, I consider two inflation indicators: (1) the log-difference of the core CPI; (2) the log- difference of the core PPI.

To capture risk-free interest rates I use: seven interest rates which include the entire term structure of interest rates. In particular, the effective federal funds rate, and constant maturity Treasury yields at 6 months, 1 year, 2 years, 3 years, 5 years, and 10 years. To obtain the approximate stationarity, I convert the nominal yields into real terms.[2]

[1]The leading index predicts the six-month growth rate of the coincident index. In addition to the coincident index, the models include other variables that lead the economy: housing permits (1 to 4 units), initial unemployment insurance claims, delivery times from the Institute for Supply Management (ISM) manufacturing survey, and the interest rate spread between the 10-year Treasury bond and the 3-month Treasury bill.

[2]Following Gilchrist et al. (2009), I employ both the realized inflation and survey measures of inflation expectations to convert the nominal yields to real terms. The realized inflation is defined as the difference between the log of the core CPI price index and its value 12 months earlier. The inflation expectation, collected by the Survey of Professional Forecasters (SPF), is at a quarterly frequency and includes the 1-year and 10-years ahead expected CPI inflation. Thus, I obtain the monthly estimates of inflation expectations from a linear interpolation of quarterly inflation expectations. Moreover, the expected inflation for 2 years, 3 years and 5 years is calculated by using the weighted average of 1-year ahead and 10-year ahead expected inflation. Simply, for calculating the 3-year expected inflation, I use the weights of 0.7 and 0.3 respectively on 1-year ahead and 10-year ahead expected inflation. Then, the real federal funds rate is measured as the difference between the nominal rate and realized inflation. The real 6-month Treasury yield is measured as the difference between the nominal yield and the equally weighted average of the realized inflation and the one-year ahead expected inflation. Remaining real Treasury yields, are obtained by the difference between the nominal yields and the appropriate expected inflation rate discussed above.

Finally, to capture financial market activity I use: (1) the total value-weighted excess market return; (2) the implied volatility on the $S\&P500$ index options (VIX) to capture uncertainty in the equity market; (3) the difference between Moody's Baa and Aaa rated corporate bond yields to capture aggregate default risk in the economy and expected excess returns on stocks and bonds; (4) the implied volatilities on Eurodollar interest rate, a measure of uncertainty of movement in short-term interest rates; (5) the implied volatilities on 10-year Treasury note futures, a measure of uncertainty of movements in long-term interest rates; (6) the log-difference of the trade-weighted exchange value of the dollar against major currencies; and (7) the difference between the 5-year swap rate and the yield on the 5-year Treasury note, a measure of liquidity.

How do I measure speculative activity? As I explained in section 1.3, the futures risk premium can be modeled in terms of the sum of the excess market return and the futures excess return. In this case, the above fundamental variables summarizing the real economy capture both excess market returns and some portion of futures excess returns which respond to fundamentals. Speculative activity is explained by the remaining part of the futures excess return.

The vector of X_{spec} contains 11 variables which reflect the information about speculators' activity in the economy. They are: (1) the variation in the aggregate level of fundamental hedging demand for oil, which can be proxied for by aggregate commodity producer default risk (EDF); (2) the realized variance of oil futures returns (RV) as a variable for time variation in the quantity of risk in the commodity market; (3) the implied crude oil volatility index (OVX), a measure of quantity of risk in the commodity market; (4) the interaction between expected default frequency of oil companies and the variance of oil futures returns ($EDF \times RV$); (5) the interaction between expected default frequency of oil companies and the implied crude oil volatility index ($EDF \times OVX$); (6) the variation in the aggregate level of fundamental hedging demand, which can be proxied for by the Zmijewski-score (Zm_score) by using the firm-level balance-sheet variables; (7) the interaction between the Zmijewski-score of oil companies and the variance of oil futures returns ($Zm\_score \times RV$); (8) the interaction between the Zmijewski-score of oil companies and the implied crude oil volatility index ($Zm\_score \times OVX$); (9) the risk tolerance of speculators, which can be proxied for by growth in intermediaries' assets relative to household asset growth as a measure of speculators' ease of access to capital in aggregate (γ_s^{-1}); (10) the interaction between expected default frequency of oil companies and the risk tolerance of speculators ($EDF \times \gamma_s^{-1}$); (11) the interaction between the Zmijewski-score of oil companies and the risk tolerance of speculators ($Zm\_score \times \gamma_s^{-1}$).

Thus in the above baseline specification, the vector X_{fund} contains 24 macroeconomic and financial time-series. The 11 elements of vector X_{spec} correspond to the variables that can summarize the speculative activities in the economy. Recalling my procedure, I first find the common factors of X_{fund}, and then extract the part of X_{spec} that is orthogonal to X_{fund}. So it is only speculative activity unrelated to fundamentals that I call "excess speculation".

In addition to the baseline specification, I consider two other specifications, including:

Specification 1: I Incorporate global economic activity by including two proxies of global economic activity in the vector $X_{fund,t}$. Following Juvenal and Petrella (2011), the first proxy which I consider is world industrial production and the second proxy is dry cargo bulk freight rates as proposed by Kilian (2009).

Specification 2: I Augment specification 1 with world oil supply. In fact, I incorporate two proxies of global economic activity as well as world oil supply in the vector $X_{fund,t}$.

2.4 Data

The data spans the period 1986-2013 for most of series, while the data for some series is available only for a subset of the period. The data sources are mentioned in section 1.5 and Appendix A. In addition, I construct proxies for some of the variables described in Appendix B. The data and proxies for variables have been tested for stationarity and those which are integrated of order one ($I(1)$) have been transformed to difference out unit roots and trends. Then, as is standard in using principal component models, the data have been standardized to have mean zero and unit standard deviation.

2.5 Empirical Results

The empirical results are shown in Table 2.1-2.5. It includes the results for different specification discussed in section 2.3 and for different sample periods based on the data availability. To study how oil risk premia have evolved over time in response to changes in fundamental and speculative activities, I also consider two distinct periods, before 2003 (subsample 1) and after 2003 (subsample 2). This can then show a possible structural break in the contribution of fundamentals and speculative activities to oil risk premia.

Table 2.1 shows the estimated contribution of fundamentals and excess speculation factors to oil risk premia in Baseline specification, including 19 fundamental variables available for the whole sample period 1986-2013. The contribution of fundamentals is about 2.4% for the whole period, 0.54% for the subsample 1, and 7.1% for the subsample 2. The contribution of second lagged of excess speculation factors to oil risk premia is about 2.5% for the whole sample, 1.3% for the subsample 1, and 16.7% for the subsample 2. The results show an increase in contribution of fundamental factors and a surge in contribution of excess speculation factors to oil risk premia after 2003.

Table 2.2 shows the estimated contribution of fundamentals and excess speculation factors to oil risk premia in Specification 1, augmenting two proxies of global economic activity to fundamental variables in Baseline specification. The sample period changes to 1994-2013 due to availability of data on world industrial production. The contribution of fundamentals is about 8.1% for the whole period, 6% for the subsample 1, and 8.4% for the subsample 2. The contribution of second lagged of excess speculation factors is about 5.5% for the whole sample, 3.1% for the subsample 1, and 13.1% for the subsample 2. The results show an increase in contribution of fundamental factors and excess speculation factors to oil risk premia after 2003.

Table 2.3 shows the estimated contribution of fundamentals and excess speculation factors to oil risk premia in Specification 2, augmenting world oil supply to fundamental variables in Specification 1. The contribution of fundamentals is about 7.9% for the whole period, 2.1% for the subsample 1 (1994-2003), and 9% for the subsample 2 (2004-2013). The contribution of second lagged of excess speculation factors is about 5.6% for the whole sample, 3.4% for the subsample 1, and 14.7% for the subsample 2. The results show an increase in contribution of fundamental factors and excess speculation factors to oil risk premia after 2003.

Table 2.4 shows the estimated contribution of fundamentals and excess speculation factors to oil risk premia in Baseline specification, including 22 fundamental variables.

The sample starts from 2004:10 since the data on three of the fundamental variables including Real personal consumption expenditures, VIX, and a measure of liquidity (defined as the difference between the 5-year swap rate and the yield on the 5-year Treasury note) is available after 2004:10. The contribution of fundamentals and first lagged of excess speculation factors to oil risk premia is about 13.1% and 19.5% respectively.

Table 2.5 shows the estimated contribution of fundamentals and excess speculation factors to oil risk premia in Baseline specification, including all 24 fundamental variables and 11 speculative variables mentioned in the empirical methodology section. The data on all fundamental variables and speculative variables, mentioned in the empirical methodology section, are available after 2008:1. The contribution of fundamentals is about 4.2% and the contribution of contemporaneous, first lagged , and second lagged excess speculation factors to oil risk premia is about 39.7%, 22.9%, and 20.1% respectively.

The results in all tables lay out a structural break in the contribution of fundamentals and speculative activities to oil risk premia after 2003, and an increased contribution of speculators in the oil market after 2003 compared to before 2003.

2.6 Conclusion

The results in all of the specifications, shown in Table 2.1-2.5, provide evidence of excess speculative activities (influence of financial investors) in oil futures markets after 2003 based on time-varying risk premia models. The results can be taken as an indicator of a structural change in oil futures markets in recent years.

Bibliography

[1] Amenguel, D., Watson, M.W., 2007. "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel," *Journal of Business and Economic Statistics*,91-96.

[2] Bai, J., Ng, S., 2002. "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70: 191-221.

[3] Gilchrist, S., Yankov, V., Zakrajsek, E., 2009. "Credit Market Shocks and Economic Fuctuations: Evidence from Corporate Bond and Stock Markets," *Journal of Monetary Economics*, 56: 471-493.

[4] Juvenal, L., Petrella, L., 2011. "Speculation in the Oil Market," Working paper. Federal Reserve Bank of St. Louis.

[5] Kilian, L., 2009. "Not all Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 99(3): 1053-1069.

[6] Zmijewski, M.E., 1984. "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," *Journal of Accounting Research*, 22: 59-82.

Table 2.1: Baseline specification. Estimated contribution of fundamentals and excess speculation factors to oil futures returns, including 19 fundamental variables available for the whole sample.

Dependent variable: oil futures returns	1986-2013	1986-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	2.36%	0.54%	7.05%
1-month ahead forecasted fundamental factors	2.59%	0.54%	6.96%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	3.27%	0.23%	0.37%
First lagged excess speculation factors	1.99%	0.15%	6.88%
Second lagged excess speculation factors	2.52%	1.31%	16.71%
Number of static fundamental factors (r_f)	2	2	2
Number of dynamic fundamental factors (q_f)	2	1	4
Number of static excess speculation factors (r_s)	3	1	1

Notes: Contributions are the share of variance of oil futures returns accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(OilFuturesReturn_t)}$ and $\frac{\beta_s^2 var(F_{exspec})}{var(OilFuturesReturn_t)}$). The estimation includes 19 fundamental variables available for the whole sample. r_f shows the number of "fundamental" factors (F_{fund}) which span all the information contained in the observed fundamental variables (X_{fund}). r_s represents the number of excess speculation factors (F_{exspec}) which span the part of information contained in the observed speculative variables (X_{spec}) unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors (q_f) shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of r_f fundamental factors (F_{fund}) at time $t+1$ as of time t is found by estimating the VAR(q_f) model for the estimated r_f static fundamental factors F_{fund}.

Table 2.2: Specification 1. Estimated contribution of fundamentals and excess speculation factors to oil futures returns, augmenting two proxies of global economic activity to fundamental variables in Baseline specification.

Dependent variable: oil futures returns	1994-2013	1994-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	8.11%	6.00%	8.42%
1-month ahead forecasted fundamental factors	9.18%	10.33%	9.57%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	3.80%	0.80%	15.84%
First lagged excess speculation factors	2.52%	0.30%	10.43%
Second lagged excess speculation factors	5.46%	3.05%	13.08%
Number of static fundamental factors (r_f)	4	4	2
Number of dynamic fundamental factors (q_f)	4	5	5
Number of static excess speculation factors (r_s)	3	1	3

Notes: Contributions are the share of variance of oil futures returns accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(OilFuturesReturn_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(OilFuturesReturn_t)}$*). The sample starts from* 1994 : 10 *due to availability of data on world industrial production.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund})*.* r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR(*q_f*) model for the estimated* r_f *static fundamental factors* F_{fund}*.*

Table 2.3: Specification 2. Estimated contribution of fundamentals and excess speculation factors to oil futures returns, augmenting world oil supply to fundamental variables in Specification 1.

Dependent variable: oil futures returns	1994-2013	1994-2003	2004-2013
Contribution of fundamentals (F_{fund})			
Contemporaneous fundamental factors	7.90%	2.13%	9.02%
1-month ahead forecasted fundamental factors	9.78%	4.93%	11.08%
Contribution of excess speculation factors (F_{exspec})			
Contemporaneous excess speculation factors	3.66%	0.30%	14.95%
First lagged excess speculation factors	2.76%	0.09%	9.53%
Second lagged excess speculation factors	5.63%	3.40%	14.71%
Number of static fundamental factors (r_f)	4	3	3
Number of dynamic fundamental factors (q_f)	4	5	4
Number of static excess speculation factors (r_s)	3	1	3

Notes: Contributions are the share of variance of oil futures returns accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(OilFuturesReturn_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(OilFuturesReturn_t)}$*). The sample starts from* 1994 : 10 *due to availability of data on world industrial production.* r_f *shows the number of "fundamental" factors* (F_{fund}) *which span all the information contained in the observed fundamental variables* (X_{fund}). r_s *represents the number of excess speculation factors* (F_{exspec}) *which span the part of information contained in the observed speculative variables* (X_{spec}) *unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors* (q_f) *shows the dynamic process between the static fundamental factors by determining the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors* (F_{fund}) *at time* $t+1$ *as of time* t *is found by estimating the VAR(*q_f*) model for the estimated* r_f *static fundamental factors* F_{fund}.

Table 2.4: Estimated contribution of fundamentals and excess speculation factors to oil futures returns in Baseline specification, including 22 fundamental variables.

Dependent variable: oil futures returns	2004-2013
Contribution of fundamentals (F_{fund})	
Contemporaneous fundamental factors	13.08%
1-month ahead forecasted fundamental factors	14.08%
Contribution of excess speculation factors (F_{exspec})	
Contemporaneous excess speculation factors	12.34%
First lagged excess speculation factors	19.50%
Second lagged excess speculation factors	10.98%
Number of static fundamental factors (r_f)	3
Number of dynamic fundamental factors (q_f)	5
Number of static excess speculation factors (r_s)	3

Notes: Contributions are the share of variance of oil futures returns accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(OilFuturesReturn_t)}$ *and* $\frac{\beta_s^2 var(F_{exspec})}{var(OilFuturesReturn_t)}$*). The data on all three fundamental variables including Real personal consumption expenditures,* VIX*, and a measure of liquidity (defined as the difference between the 5-year swap rate and the yield on the 5-year Treasury note) are available after* 2004 : 10*. The results are based on Baseline specification including 22 fundamental variables.* r_f *shows the number of "fundamental" factors (*F_{fund}*) which span all the information contained in the observed fundamental variables (*X_{fund}*).* r_s *represents the number of excess speculation factors (*F_{exspec}*) which span the part of information contained in the observed speculative variables (*X_{spec}*) unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors (*q_f*) determines the order of autoregressive process for static fundamental factors. 1-month ahead expectation of* r_f *fundamental factors (*F_{fund}*) at time* $t+1$ *as of time* t *is found by estimating the VAR(*q_f*) model for the estimated* r_f *static fundamental factors* F_{fund}*.*

Table 2.5: Estimated contribution of fundamentals and excess speculation factors to oil futures returns in Baseline specification, including all 24 fundamental variables and 11 speculative variables mentioned in the empirical methodology section.

Dependent variable: oil futures returns	2008-2013
Contribution of fundamentals (F_{fund})	
Contemporaneous fundamental factors	4.19%
1-month ahead forecasted fundamental factors	4.19%
Contribution of excess speculation factors (F_{exspec})	
Contemporaneous excess speculation factors	39.65%
First lagged excess speculation factors	22.88%
Second lagged excess speculation factors	20.14%
Number of static fundamental factors (r_f)	2
Number of dynamic fundamental factors (q_f)	1
Number of static excess speculation factors (r_s)	4

Notes: Contributions are the share of variance of oil futures returns accounted for by the latent factors (as measured by $\frac{\beta_f^2 var(F_{fund})}{var(OilFuturesReturn_t)}$ and $\frac{\beta_s^2 var(F_{exspec})}{var(OilFuturesReturn_t)}$). The data on all 24 fundamental variables and 11 speculative variables, mentioned in the empirical methodology section, are available after 2008 : 1. The results are based on Baseline specification including 24 fundamental variables and 11 speculative variables. r_f shows the number of "fundamental" factors (F_{fund}) which span all the information contained in the observed fundamental variables (X_{fund}). r_s represents the number of excess speculation factors (F_{exspec}) which span the part of information contained in the observed speculative variables (X_{spec}) unrelated to fundamentals and specifically related to speculative activities. The number of dynamic fundamental factors (q_f) determines the order of autoregressive process for static fundamental factors. 1-month ahead expectation of r_f fundamental factors (F_{fund}) at time $t+1$ as of time t is found by estimating the VAR(q_f) model for the estimated r_f static fundamental factors F_{fund}.